

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	22.4	100.0	28	6	CQ898516	CQ898516 Trypanoso
2	18	80.4	148802	2	AC159703	AC159703 Continuation (24 o
3	17.8	79.5	110000	1	AP006618_23	Continuation (30 o
4	17.8	79.5	110000	1	BA00035_29	Continuation (30 o
5	17.6	78.6	1149	6	AR671202	AR671202 Sequence
6	17.6	78.6	110000	1	BR000112_68	Continuation (69 o
7	17.6	78.6	148213	9	AL645731	AL645731 Mouse DNA
8	17.6	78.6	227259	14	AC120702	AC120702 Rattus no
9	17.6	78.6	238932	14	AC096005	AC096005 Rattus no
10	17.6	78.6	241954	14	AC126116	AC126116 Rattus no
11	17.4	77.7	101783	14	AP007942	AP007942 Lotus cor
12	17.4	77.7	156795	14	AC135389	AC135389 Rattus no
13	17.4	77.7	169506	9	AC123818	AC123818 Mus muscu
14	17.4	77.7	185530	14	AC109026	AC109026 Rattus no
15	17.2	76.8	110000	1	AP006840_33	Continuation (34 o
16	17.2	76.8	116522	8	AL445228	AL445228 Human DNA
17	17.2	76.8	157078	14	AC121229	AC121229 Rattus no
18	17.2	76.8	175413	14	AC147980	AC147980 Rattus no

AC159703
 AC159703.1 GI:62359576
 HTG.
 SOURCE Trypanosoma brucei
 ORGANISM Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

REFERENCE
 AUTHORS
 1 (bases 1 to 148802)
 Ghedin, E., Blandin, G., Bartholomeu, D., Caler, E., Haas, B., Hannick, L., Shallom, J., Hou, L., Djikeng, A., Feldblyum, T., Hostetler, J., Johnson, J., Jones, K., Koo, H. L., Larkin, C., Pal, G., Peterson, J., Khalak, H. G., Salzberg, S., Simpson, A. J., Tallon, L., Van Aken, S., Wanless, D., White, O., Wortman, J., Fraser, C. M. and El-Sayed, N. M. A.
 Unpublished

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 2 (bases 1 to 148802)
 Haas, B., Blandin, G. and El-Sayed, N.
 Direct Submission
 Submitted (07-APR-2005) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, nelsayed@tigr.org
 BAC clone RPC193-30K1 from Trypanosoma brucei chromosome 8.
 Putative protein coding regions (>30 codons) were predicted by combining the output of the gene finding algorithm GIMMER with similarity data and manual annotation of open reading frames.
 Further refinement of gene models (additions, deletions and alterations to boundary coordinates) were done manually based on detailed comparisons with the T. cruzi and L. major genomes. Genes on BACs were assigned a systematic name based on the chromosome and the BAC from which they originated (e.g. Tb03.27P10.410). Gene products were assigned a putative function when they shared significant similarity with experimentally characterized gene products or when they contained functionally known protein domains. Gene products were labelled either 'hypothetical protein', or 'hypothetical protein, conserved' in the case of products showing significant similarity with proteins or domains of unknown function from other organisms. Gene products of unknown function with predicted orthologs in Trypanosoma cruzi and Leishmania major genomes were automatically classified as conserved proteins. Short predicted coding sequences (less than 150 codons) with no additional evidence of coding potential were labelled 'hypothetical protein, unlikely'.

FEATURES
 source
 1..148802
 /organism="Trypanosoma brucei"
 /mol_type="genomic DNA"
 /isolate="GUTat10.1"
 /db_xref="taxon:5691"
 /chromosome="8"
 /clone="RPC193-30K1"
 1..21371
 /note="annotation for this sequence region can be found on overlapping clone 11J15."

misc_feature
 567..587
 /note="AT rich"

misc_feature
 1052..1074
 /note="microsatellite"

repeat_region
 /rpt_type=tandem
 /rpt_unit="ta"
 3341..3381
 /note="AT rich"

misc_feature
 3563..3649
 /note="T-rich"

repeat_region
 4450..4475
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ta"

repeat_region
 4497..4516
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ggaaa"

misc_feature
 7029..7079
 /note="GA-rich"

repeat_region
 12175..12208

/note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ttg"
 13090..13111
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ta"
 13233..13256
 /note="AT rich"
 13356..13375
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ttg"
 13417..13461
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ta"
 13766..13812
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="caa"
 14376..14409
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="caa"
 14475..14502
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ttttc"
 14627..14654
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="taa"
 14717..14743
 /note="AT rich"
 14872..14901
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="gaaa"
 15099..15127
 /note="AT rich"
 15448..15492
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="cca"
 15540..15617
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="gaatg"
 15738..15762
 /note="AT rich"
 15985..16007
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="gaaa"
 16762..16811
 /note="CT-rich"
 16837..16858
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ca"
 17325..17381
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ttta"
 19235..19254
 /note="microsatellite"
 /rpt_type=tandem
 /rpt_unit="ta"
 <21372..>22322
 /locus_tag="Tb927.8.6480"
 /note="synonyms: Tb08.30K1.220, Tb08.11J15.05; hypothetical protein, conserved"

WPCOMMENT
Sequence split into 71 fragments LOCUS BA000012 Accession BA000012

Fragment Name	Begin	End
BA000012_00	1	110000
BA000012_01	100001	210000
BA000012_02	200001	310000
BA000012_03	300001	410000
BA000012_04	400001	510000
BA000012_05	500001	610000
BA000012_06	600001	710000
BA000012_07	700001	810000
BA000012_08	800001	910000
BA000012_09	900001	1010000
BA000012_10	1000001	1110000
BA000012_11	1100001	1210000
BA000012_12	1200001	1310000
BA000012_13	1300001	1410000
BA000012_14	1400001	1510000
BA000012_15	1500001	1610000
BA000012_16	1600001	1710000
BA000012_17	1700001	1810000
BA000012_18	1800001	1910000
BA000012_19	1900001	2010000
BA000012_20	2000001	2110000
BA000012_21	2100001	2210000
BA000012_22	2200001	2310000
BA000012_23	2300001	2410000
BA000012_24	2400001	2510000
BA000012_25	2500001	2610000
BA000012_26	2600001	2710000
BA000012_27	2700001	2810000
BA000012_28	2800001	2910000
BA000012_29	2900001	3010000
BA000012_30	3000001	3110000
BA000012_31	3100001	3210000
BA000012_32	3200001	3310000
BA000012_33	3300001	3410000
BA000012_34	3400001	3510000
BA000012_35	3500001	3610000
BA000012_36	3600001	3710000
BA000012_37	3700001	3810000
BA000012_38	3800001	3910000
BA000012_39	3900001	4010000
BA000012_40	4000001	4110000
BA000012_41	4100001	4210000
BA000012_42	4200001	4310000
BA000012_43	4300001	4410000
BA000012_44	4400001	4510000
BA000012_45	4500001	4610000
BA000012_46	4600001	4710000
BA000012_47	4700001	4810000
BA000012_48	4800001	4910000
BA000012_49	4900001	5010000
BA000012_50	5000001	5110000
BA000012_51	5100001	5210000
BA000012_52	5200001	5310000
BA000012_53	5300001	5410000
BA000012_54	5400001	5510000
BA000012_55	5500001	5610000
BA000012_56	5600001	5710000
BA000012_57	5700001	5810000
BA000012_58	5800001	5910000
BA000012_59	5900001	6010000
BA000012_60	6000001	6110000
BA000012_61	6100001	6210000
BA000012_62	6200001	6310000
BA000012_63	6300001	6410000
BA000012_64	6400001	6510000
BA000012_65	6500001	6610000
BA000012_66	6600001	6710000
BA000012_67	6700001	6810000
BA000012_68	6800001	6910000
BA000012_69	6900001	7010000

Continuation (30 of 32) of BA000035 from base 2900001 (BA000035 Corynebacterium efficiens)

Query Match 79.5%; Score 17.8; DB 1; Length 110000;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 15; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CGGAUGCGGUCDCDADGDTGD 27
|||:||||:||||:||||:||||:||||:
Db 95409 CGGATGCGGTCTTCGCAAGCTAGA 95385

RESULT 5
AR671202/c AR671202 1149 bp DNA linear PAT 13-JUN-2005
LOCUS Sequence 703 from patent US 6902887.
DEFINITION AR671202
ACCESSION AR671202
VERSION AR671202.1 GI:67610848
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1149)
AUTHORS Berka,R.M., Rey,M.W., Shuster,J.R., Kauppinen,S., Clausen,I.G. and Olsen,P.B.
TITLE Methods for monitoring multiple gene expression
JOURNAL Patent: US 6902887-A 703 07-JUN-2005;
Novozymes Biotech, Inc. and Novozymes A/S; Davis, CA
FEATURES
source 1..1149
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 78.6%; Score 17.6; DB 6; Length 1149;
Best Local Similarity 53.6%; Pred. No. 1.9e+03;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGAUGCGGUCDCDADGDTGD 28
|||:||||:||||:||||:||||:||||:
Db 664 GATGATGCGGTCTTCGAGCATTTGGT 637

RESULT 6
BA000012_68/c

BA000012.70 7000001 7036071
Continuation (69 of 71) of BA000012 from base 6800001 (BA000012 Mesorhizobium loti MAFF3)

Query Match 78.6%; Score 17.6; DB 1; Length 110000;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCDCDADGDTGDT 28
DB 53329 GACGATCGGCTGCGGATGTGAT 53302

RESULT 7
AL645731
LOCUS Mouse DNA sequence from clone RP23-423G13 on chromosome 4, complete
DEFINITION AL645731 148213 bp DNA linear ROD 09-AUG-2002
ACCESSION AL645731
VERSION AL645731.8 GI:22204303
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 148213)

REFERENCE
AUTHORS Direct Submission
TITLE Submitted (01-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
COMMENT humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Aug 11, 2002 this sequence version replaced gi:21738431.

Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw.,
SWISSPROT; Tr., TrEMBL; Wp, WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-423G13 is
from the RP23-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6.

FEATURES
source
1. 148213
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone_lib="RP23-423G13"
/clone_lib="RPCI-23"

ORIGIN
Query Match 78.6%; Score 17.6; DB 9; Length 148213;
Best Local Similarity 53.6%; Pred. No. 1.5e+03;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCDCDADGDTGDT 28
DB 56798 GACGATCGGCTCTCTCTATGTAAGAT 56825

RESULT 8
AC120702/c

LOCUS Rattus norvegicus clone CH230-145L2, *** SEQUENCING IN PROGRESS
DEFINITION AC120702 227259 bp DNA linear HTG 09-OCT-2002
*** 3 unordered pieces.

ACCESSION AC120702.5 GI:23195861
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 227259)

REFERENCE
AUTHORS Muzny, D. Marie, Metzker, M. Lee, Abrazon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P.,
Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Y., Chu, J.,
Cleaveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Dengson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gubaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuwar, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poldexter, A., Popovic, D., Primus, E., Pul, L.,
Puzo, M., Quiros, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Snajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Vallas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, R., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished
2 (bases 1 to 227259)
REFERENCE
AUTHORS Worley, K.C.

Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 241954)
Worley,K.C.

Direct Submission
Submitted (03-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 241954)

Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24635886.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GABP
Center clone name: CH230-3D21
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 219177 bases at least Q40
Consensus quality: 22082 bases at least Q30
Consensus quality: 224351 bases at least Q20
Estimated insert size: 230914; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 240470: contig of 240470 bp in length
* 240471 240570: gap of unknown length
* 240571 241954: contig of 1384 bp in length.

FEATURES
source
1..241954
/organism="Rattus norvegicus"
/mol_type="Genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3D21"
1..2456
misc_feature
/note="wgs contig"
240471..240570
gap
/estimated_length=unknown

ORIGIN
Query Match 78.6%; Score 17.6; DB 14; Length 241954;
Best Local Similarity 53.6%; Pred. No. 1.5e+03;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
QY 1 GACGCGGCGGUCUDCCDADGDTGDT 28
Db 176301 GATGGTGGGCTGCACTAGTAGGT 176274

RESULT 11
AP007942/c
LOCUS
DEFINITION
AP007942 101783 bp DNA linear HTG 28-DEC-2004
SEQUENCING IN PROGRESS ***, 17 unordered pieces.
ACCESSION
AP007942 GI:56806250
VERSION
HTG; HTGS PHASE1.
KEYWORDS
Lotus corniculatus var. japonicus (Lotus japonicus)
SOURCE
Lotus corniculatus var. japonicus
ORGANISM
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.

REFERENCE
1
AUTHORS
Kaneko,T., Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE
Structural Analysis of a Lotus Japonicus Genome. XI. Sequence Features and Mapping of Nine hundred twenty-one TAC Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 101783)
AUTHORS
Sato,S.
TITLE
Direct Submission
JOURNAL
Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp, URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337), Fax:81-438-52-3934)
* NOTE: this is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 7386: contig of 7386 bp in length
* 7387 7486: gap of unknown length
* 7487 8888: contig of 1402 bp in length
* 8889 8988: gap of unknown length
* 8989 10283: contig of 1295 bp in length
* 10284 10383: gap of unknown length
* 10384 12992: contig of 2609 bp in length
* 12993 13092: gap of unknown length
* 13093 15958: contig of 2866 bp in length
* 15959 16058: gap of unknown length
* 16059 19435: contig of 3377 bp in length
* 19436 19535: gap of unknown length
* 19536 22371: contig of 2836 bp in length
* 22372 22471: gap of unknown length
* 22472 25286: contig of 2815 bp in length
* 25287 25386: gap of unknown length
* 25387 29278: contig of 3892 bp in length
* 29279 29378: gap of unknown length
* 29379 36932: contig of 7554 bp in length
* 36933 37032: gap of unknown length
* 37033 42634: contig of 5602 bp in length
* 42635 42734: gap of unknown length
* 42735 51642: contig of 8908 bp in length
* 51643 51742: gap of unknown length
* 51743 57986: contig of 6244 bp in length
* 57987 58086: gap of unknown length
* 58087 67354: contig of 9268 bp in length
* 58087


```

* 67355 67454: gap of unknown length
* 67455 79586: contig of 12132 bp in length
* 79587 79686: gap of unknown length
* 79687 101083: contig of 21397 bp in length
* 101084 101183: gap of unknown length
* 101184 101783: contig of 600 bp in length.
Location/Qualifiers
1. 101783
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variant="japonicus"
/variant="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone_lib="LJT library"
/clone="LJT31111"
/notes="PAC clone: TM1265, synonym: Lotus japonicus"
7387. 7486
/estimated_length=unknown
8889. 8988
/estimated_length=unknown
10284. 10383
/estimated_length=unknown
12993. 13092
/estimated_length=unknown
15959. 16058
/estimated_length=unknown
19436. 19535
/estimated_length=unknown
22372. 22471
/estimated_length=unknown
25287. 25386
/estimated_length=unknown
29279. 29378
/estimated_length=unknown
36933. 37032
/estimated_length=unknown
42635. 42734
/estimated_length=unknown
51643. 51742
/estimated_length=unknown
57987. 58086
/estimated_length=unknown
67355. 67454
/estimated_length=unknown
79587. 79686
/estimated_length=unknown
101084. 101183
/estimated_length=unknown

```

ORIGIN

```

Query Match 77.7%; Score 17.4; DB 14; Length 101783;
Best Local Similarity 55.6%; Pred. No. 1.9e+03;
Matches 15; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Oy 2 ACAGGATGAGTTCGCGAAGCTGTT 44997
||||| :|: :|: :|: :|: :|:
Db 44923 ACAGGATGAGTTCGCGAAGCTGTT 44997

RESULT 12
AC135389/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-477F9, *** SEQUENCING IN PROGRESS
***
AC135389
AC135389.2 GI:25007661
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 156795)

```

AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alibabek, S., Amin, A., Angiano, D., Anylebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, P., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoggues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenschew, L., Louised, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Snares, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

Unpublished

2 (bases 1 to 156795)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (14-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 156795)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 15, 2002 this sequence version replaced gi:23955973. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

SOURCE	ORGANISM
Mus musculus (house mouse)	Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.	1 (bases 1 to 169506)
Cedroni, M. and Kozlowski, A.	The sequence of Mus musculus BAC clone RP24-266A12
Unpublished (2001)	Unpublished (2001)

```

70678. .70747
/product="tRNA-Thr"
/note="Likely_pseudogen
Sc=-.9.72"
repeat_region
/rpt_family="t1"
74429. 74520
/rpt_family="t1"
85371. 85498
/rpt_family="Alu"
repeat_region
90698. 90828
/rpt_family="Alu"
repeat_region
99035. 99445
/rpt_family="MaLR"
99809. 99916
/rpt_family="MER2_type"
102750. 102958
/rpt_family="B2"
repeat_region
102983. 103111
/rpt_family="B4"
repeat_region
105264. 105347
/rpt_family="t1"
repeat_region
105354. 105489
/rpt_family="Alu"
repeat_region
105591. 105857
/rpt_family="B4"
repeat_region
109201. 109417
/rpt_family="B4"
repeat_region
112035. 112306
/rpt_family="B4"
repeat_region
113495. 113625
/rpt_family="t1"
repeat_region
113683. 113828
/rpt_family="MaLR"
113882. 113995
/rpt_family="MaLR"
repeat_region
119891. 120110
/rpt_family="B2"
repeat_region
120821. 121034
/rpt_family="B4"
repeat_region
125617. 126698
misc feature

```

Q7 1 GACGAUGCGGUCUDCDADGTDGD 27
 ||||| : | : : : | : : :
D6 138265 GACGATGTGATCTACACCAAGTTGT 138239

RESULT 14					
AC109026/c					
LOCUS	AC109026	185530 bp	DNA	linear	HTG 11-OCT-2002
DEFINITION	Rattus norvegicus clone CH230-27907.	***	SEQUENCING IN PROGRESS		

ACCESSION	AC109026
VERSION	AC109026.4
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS

1. (bases 1 to 195530)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Aisbrooks, S., Amin, A., Anguitano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, S., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Celler project name: GERM

AP006840_21 2100001 2210000
 AP006840_22 2200001 2310000
 AP006840_23 2300001 2410000
 AP006840_24 2400001 2510000
 AP006840_25 2500001 2610000
 AP006840_26 2600001 2710000
 AP006840_27 2700001 2810000
 AP006840_28 2800001 2910000
 AP006840_29 2900001 3010000
 AP006840_30 3000001 3110000
 AP006840_31 3100001 3210000
 AP006840_32 3200001 3310000
 AP006840_33 3300001 3410000
 AP006840_34 3400001 3510000
 AP006840_35 3500001 3566135

Continuation (34 of 36) of AP006840 from base 3300001 (AP006840 Symbiobacterium thermophilus)

Query Match 76.8%; Score 17.2; DB 1; Length 110000;
 Best Local Similarity 57.7%; Pred. No. 2.4e+03;
 Matches 15; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGGAUGCGGUCUCCDADGDTG 26
 Db 86965 GACGTTGGGTCCTCTGATCATGG 86990

Search completed: February 18, 2006, 15:20:06
 Job time : 937.581 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2006, 14:29:17 ; Search time 207.29 Seconds
(without alignments)
900.242 Million cell updates/sec

Title: US-10-824-829-1

Perfect score: 22.4
Sequence: 1 gacggagcgucdcdadgdtdgdt 28

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.6	78.6	606	13 ACN46113	Cotton pr
2	17.6	78.6	1149	3 AAF08180	Fusarium
3	17.6	78.6	1149	13 ADU5221	Adus2221 Fusarium
4	17.6	78.6	1149	13 ADZ90224	Adz90224 Fusarium
5	17.2	76.8	506	13 ADR65506	Adr65506 Corton cd
6	17.2	76.8	622	10 ADG37889	Adg37889 Aspergill
7	17	75.9	720	10 ABZ41650	Abz41650 N. gonorr
8	17	75.9	720	10 ABZ37718	Abz37718 N. gonorr
9	17	75.9	765	10 ABZ41652	Abz41652 N. gonorr
10	16.6	74.1	1377	14 ACL70886	ACL70886 M. xanthu
11	16.6	74.1	5515	14 ACL64238	ACL64238 M. xanthu
12	16.6	74.1	110000	12 ADN46845_05	Continuation (6 of
13	16.6	74.1	110000	12 ADN47591_15	Continuation (16 o
14	16.6	74.1	110000	12 ADN46123_05	Continuation (6 of
15	16.6	74.1	110000	12 ADN47209_15	Continuation (16 o
16	16.6	74.1	110000	12 ADN46464_05	Continuation (6 of
17	16.6	74.1	110000	12 ADN47960_15	Continuation (16 o
18	16.2	72.3	1125	8 ACA37897	ACA37897 Prokaryot
19	16.2	72.3	1629	5 AAF31537	AAF31537 C.glutami

20	16.2	72.3	1704	5 AAH67087	AAH67087 C glutami
21	16.2	72.3	1704	14 AEB13114	AEB13114 C. glutami
22	16.2	72.3	1827	5 AAF31536	AAF31536 C. glutami
23	16.2	72.3	1827	6 ABS65349	ABS65349 DNA encod
24	16.2	72.3	1834	9 ACC59425	ACC59425 C glutami
25	16.2	72.3	2005	6 ABK49953	ABK49953 Phosphotr
26	16.2	72.3	2005	6 ABK49954	ABK49954 Phosphotr
27	16.2	72.3	2298	6 ABK49956	ABK49956 Phosphotr
28	16.2	72.3	27684	11 ADL27140	ADL27140 Human gen
29	16.2	72.3	55235	4 AAK67426	AAK67426 Human inm
30	16.2	72.3	349980	5 AAH68530	AAH68530 C glutami
31	16	71.4	603	5 ABV56449	ABV56449 Human pro
32	16	71.4	792	4 ABL29149	ABL29149 Drosophil
33	16	71.4	983	13 ADS63204	ADS63204 Bacterial
34	16	71.4	983	13 ADS62208	ADS62208 Bacterial
35	16	71.4	983	13 ADS59787	ADS59787 Bacterial
36	16	71.4	983	13 ADS62721	ADS62721 Bacterial
37	16	71.4	1062	6 ABZ12975	ABZ12975 Arabidops
38	16	71.4	1194	8 ACAS3616	ACAS3616 Prokaryot
39	16	71.4	1197	2 AAV42059	AAV42059 Kurthia s
40	16	71.4	1378	3 AAC49685	AAC49685 Arabidops
41	16	71.4	1380	3 AAC38757	AAC38757 Arabidops
42	16	71.4	2034	13 ADR85850	ADR85850 Aspergill
43	16	71.4	2184	2 AAX22318	AAX22318 B. fragil
44	16	71.4	2208	2 AAZ52917	AAZ52917 Human pro
45	16	71.4	2260	12 ADO35646	ADO35646 Novel mou

ALIGNMENTS

RESULT 1

ACN46113
ID ACN46113 standard; cDNA; 606 BP.

AC ACN46113;

DT 02-DEC-2004 (first entry)

DE Cotton primed seed EST Clone ID: LIB3825-015-Q1-K6-D10, SEQ:894.

KW Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;

KW variety DP508; library LIB3825; molecular tag; molecular marker;

KW genetic mapping; molecular mapping; seed germination; plant growth;

KW plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

XX US2004123340-A1.

PD 24-JUN-2004.

PF 12-DEC-2001; 2001US-00021323.

PR 14-DEC-2000; 2000US-0255619P.

PA (DEIK/) DEIKMAN J.

PA (FENG/) FENG P C C.

PA (FINC/) FINCHER K L.

PA (ZIEG/) ZIEGLER T E.

PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;

DR WPI; 2004-479808/45.

PT New isolated nucleic acid molecule that encodes a plant protein or its fragment, useful for isolating a variety of agronomically significant genes associated with plant growth, quality or yield, and as molecular tags to map genes.

PS Claim 1; SEQ ID NO 894; 34pp; English.

CC The invention relates to 17880 cotton expressed sequence tags (ESTs;

•

XX Monitoring differential expression of genes in a filamentous fungal cell
PI relative to expression of the genes in second filamentous fungal cells,
PT comprises using microarrays containing *Fusarium venenatum* expressed
PT sequence tags.
XX
XX Claim 1; SEQ ID NO 703; 274pp; English.
XX
CC The invention relates to a method for monitoring the differential
CC expression of genes in a first filamentous fungal cell relative to the
CC same genes in one or more second filamentous fungal cells. The method
CC involves differentially labeling nucleic acids from the first and second
CC fungal cells with fluorescent reporter dyes, and adding the mixture of
CC fluorescently labeled polynucleotides to a substrate containing an array
CC of *Fusarium venenatum* expressed sequence tags (ESTs) selected from
CC ADU51519-ADU55288 under conditions that permit hybridization. Relative
CC expression of genes in the fungal cells is determined by fluorescence;
CC the fluorescent signal emitted when labeled nucleic acids from the first
CC fungal cell bind to the ESTs in the array differs in color from that
CC emitted when labeled nucleic acids from the second fungal cell hybridize.
CC When both sets of labeled nucleic acids bind to the array, a distinct
CC combined fluorescence emission color is produced. The filamentous fungal
CC cells used in the method are chosen from *Acremonium*, *Aspergillus*,
CC *Fusarium*, *Hemicella*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*,
CC *Thielavia*, *Tolypocladium*, and *Trichoderma* cells, preferably *Fusarium*
CC *venenatum*, *Aspergillus niger*, or *Aspergillus oryzae*. The method of the
CC invention is useful for monitoring the expression of a plurality of genes
CC in filamentous fungal cells, in order to improve the cells' protein
CC production capacity when such organisms are used for the industrial
CC production of proteins (e.g., enzymes). The present sequence represents
CC an EST derived from *Fusarium venenatum* CCL-3 (a morphological mutant of
CC strain ATCC 20334) which is specifically claimed for use in the method of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=US20040222367.
XX
SQ Sequence 1149 BP; 262 A; 328 C; 257 G; 300 T; 0 U; 2 Other;
Query Match 78.6%; Score 17.6; DB 13; Length 1149;
Best Local Similarity 53.6%; Pred. No. 1.7e+02;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
RESULT 4
ADZ90224/c
ID ADZ90224 standard; cDNA; 1149 BP.
XX
AC ADZ90224;
XX
XX 11-AUG-2005 (first entry)
XX
DE *Fusarium venenatum* expressed sequence tag cDNA SEQ ID NO 703.
XX
KW differential expression; gene expression; filamentous fungus;
KW expressed sequence tag; EST; variation; microarray; ss.
XX
OS *Fusarium venenatum*.
XX
FN US6902887-B1.
XX
PD 07-JUN-2005.
XX
XX 22-MAR-2000; 2000US-00533559.
XX
XX 22-MAR-1999; 99US-00273623.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.

XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX WPI; 2005-401635/41.
XX
XX Monitoring differential expression of genes in first filamentous fungal
PT cell relative to that of genes in second fungal cells, comprises
PT hybridizing fluorescence labeled nucleic acids from the cells to an array
PT of expressed sequence tags.
XX
XX Disclosure; SEQ ID NO 703; 264pp; English.
XX
CC The invention relates to a method of monitoring (M1) differential
CC expression of multiple genes in a first filamentous fungal cell relative
CC to expression of the same genes in one or more second filamentous fungal
CC cells, comprises adding fluorescence labeled nucleic acids isolated from
CC the cells to a substrate containing an array of *Aspergillus oryzae*
CC expressed sequence tag (EST) of ADZ93898-ADZ96922, and examining the
CC array under fluorescence excitation conditions. (M1) is useful for
CC monitoring global expression of several genes from a filamentous fungal
CC cell, discovering new genes, identifying possible functions of unknown
CC open reading frames and monitoring gene copy number variation and
CC stability. In (M1), one spot on an array equals one gene or open reading
CC frame, extensive follow-up characterization is unnecessary since sequence
CC information is available and EST microarrays can be organized based on
CC function of the gene products. This sequence corresponds to an EST
CC sequence of the invention. (Note: this sequence is not given in the
CC printed specification but can be obtained in electronic form from the
CC USPTO web site seqdata.uspto.gov/sequence.html; Document ID: 6902887B1).
XX
SQ Sequence 1149 BP; 262 A; 328 C; 257 G; 300 T; 0 U; 2 Other;
Query Match 78.6%; Score 17.6; DB 14; Length 1149;
Best Local Similarity 53.6%; Pred. No. 1.7e+02;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
RESULT 5
ADR65506
ID ADR65506 standard; cDNA; 506 BP.
XX
AC ADR65506;
XX
XX 02-DEC-2004 (first entry)
XX
DE Cotton cDNA sequence, SEQ ID 6287.
XX
KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
KW drought tolerance; plant disease resistance; galactomannan; lignin;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW homologous recombination; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW stress resistance.
XX
OS *Gossypium hirsutum*.
XX
FN US2004181830-A1.
XX
PD 16-SEP-2004.
XX
XX 29-JAN-2004; 2004US-00767795.
XX
XX 07-MAY-2001; 2001US-00849529.
XX
XX 12-DEC-2001; 2001US-00021323.
XX
XX (KOVA/) KOVALIC D K.
XX (ZHOU/) ZHOU Y.
XX (CAOY/) CAO Y.
XX

PI Kovalic DK, Zhou Y, Cao Y;
 XX WPI; 2004-667718/65.
 XX
 XX New recombinant nucleic acid molecules and polypeptides from *Gossypium*
 PT hirsutum, useful for producing plants with improved biological
 PT characteristics (e.g. improved plant cold or drought tolerance).
 XX
 XX Claim 1; SEQ ID NO 6287; 14pp; English.
 XX
 XX The invention relates to a recombinant polynucleotide comprising any of
 CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
 CC Also a recombinant polypeptide comprising any of the 58798 amino acid
 CC sequences mentioned in the specification and producing a plant having an
 CC improved property. Producing a plant having an improved property
 CC comprises transforming a plant with a recombinant construct comprising a
 CC promoter region functional in a plant cell operably joined to a
 CC polynucleotide comprising a coding sequence for a polypeptide associated
 CC with the property, and growing the transformed plant. The polypeptide is
 CC useful for improving plant cold tolerance, manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, improving plant
 CC drought tolerance, providing increased resistance to plant disease,
 CC producing galactomannan (or lignin or plant growth regulators), improving
 CC plant heat tolerance, improving plant tolerance to herbicides, increasing
 CC the rate of homologous recombination in plants, improving plant tolerance
 CC to extreme osmotic conditions or to pathogens or pests, improving yield
 CC by modification of photosynthesis, modifying seed oil or protein yield
 CC and/or content, improving yield by modification of carbohydrate, nitrogen
 CC or phosphorus use and/or uptake, or improving yield by providing improved
 CC plant growth and development under at least one stress condition. The
 CC polynucleotide and polypeptide may also be used in recombinant DNA
 CC constructs, in physical arrays of molecules, as plant breeding markers,
 CC or in computer-based storage and analysis systems. The present sequence
 CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585
 CC polynucleotide sequences were available, the remaining 52213
 CC polynucleotides and all 58798 protein sequences were not present.
 XX
 XX Sequence 506 BP; 146 A; 109 C; 107 G; 144 T; 0 U; 0 Other;
 SQ

Query Match 76.8%; Score 17.2; DB 13; Length 506;
 Best Local Similarity 57.7%; Pred. No. 2.4e+02;
 Matches 15; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACAGGATCGGTCGCTCAGATTGA 339
 DB 314 ACAGGATCGGTCGCTCAGATTGA 339

RESULT 6
 ADG37889/c
 ID ADG37889 standard; DNA; 622 BP.
 XX
 XX AC ADG37889;
 XX
 XX DT 26-FEB-2004 (first entry)
 XX
 XX DE Aspergillus solid-culture DNA #600.
 XX
 XX KW ds; filamentous-fungus; Aspergillus fermentation; solid culture.
 XX
 XX OS Aspergillus oryzae.
 XX
 XX PN JP2003180365-A.
 XX
 XX PD 02-JUL-2003.
 XX
 XX PF 17-DEC-2001; 2001JP-00383645.
 XX
 XX PR 17-DEC-2001; 2001JP-00383645.

PA (DOKU-) DOKURITSU GYOSEI HOJIN SHURUI SOGO KENKY.
 XX
 XX DR WPI; 2003-819301/77.
 XX
 XX PT New DNA derived from *Aspergillus*, expressed by filamentous-fungi when
 PT cultured in a solid medium, useful for measuring *Aspergillus* fermentation
 PT conditions.
 XX
 XX PS Claim 2; SEQ ID NO 600; 302pp; Japanese.
 XX
 XX CC The invention relates to a DNA expressed by filamentous-fungi. The DNA is
 CC useful for detecting filamentous-fungi and is useful for measuring
 CC *Aspergillus* fermentation conditions. A method of detecting the DNA is
 CC useful for measuring the actual growth state of *Aspergillus*
 CC quantitatively and the expression level of desired gene specifically. The
 CC present sequence is used in the exemplification of the present invention.
 XX
 XX SQ Sequence 622 BP; 132 A; 182 C; 157 G; 151 T; 0 U; 0 Other;
 SQ

Query Match 76.8%; Score 17.2; DB 10; Length 622;
 Best Local Similarity 57.7%; Pred. No. 2.5e+02;
 Matches 15; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACCGAUGCGGUCUDCCDADGDTGDG 26
 DB 374 GACCGAUGCGGTCGCTCAGCAAGTTTG 349

RESULT 7
 ABZ41650/c
 ID ABZ41650 standard; DNA; 720 BP.
 XX
 XX AC ABZ41650;
 XX
 XX DT 07-MAR-2003 (first entry)
 XX
 XX DE N. gonorrhoeae nucleotide sequence SEQ ID 7889.
 XX
 XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 XX OS *Neisseria gonorrhoeae*.
 XX
 XX PN WO200279243-A2.
 XX
 XX PD 10-OCT-2002.
 XX
 XX PF 12-FEB-2002; 2002WO-IB002069.
 XX
 XX PR 12-FEB-2001; 2001GB-00003424.
 XX
 XX PA (CHIR-) CHIRON SPA.
 XX
 XX PI Fontana MR, Pizza M, Massignani V, Monaci E;
 XX
 XX DR WPI; 2003-058415/05.
 XX
 XX DR P-PSDB; ABP80680.
 XX
 XX PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX
 XX PS Disclosure; Page 763; 815pp; English.
 XX
 XX CC The present invention relates to proteins from *Neisseria gonorrhoeae*.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
 CC molecules of the invention
 XX
 XX SQ Sequence 720 BP; 155 A; 245 C; 199 G; 121 T; 0 U; 0 Other;

Query Match: 75.9%; Score 17; DB 10; Length 720;
Best Local Similarity 60.0%; Pred. No. 3.2e+02;
Matches 15; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 3 CGGAUGCGGUCDCDADGDTGDD 27
DB 88 CGGATCGGTCAGCAGCGGTAGA 64

RESULT 8
ID AB237718/c
XX AB237718 standard; DNA; 720 BP.
AC AB237718;
DT 07-MAR-2003 (first entry)
DE N. gonorrhoeae nucleotide sequence SEQ ID 25.
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX Neisseria gonorrhoeae.
XX WO200279243-A2.
XX 10-OCT-2002.
XX 12-FEB-2002; 2002WO-IB002069.
XX 12-FEB-2001; 2001GB-00003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Massignani V, Monaci E;
DR WPI; 2003-058415/05.
DR P-PSDB; ABP76748.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
medicament for treating or preventing N. gonorrhoeae infection.
PS Claim 6; Page 175; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records AB237706-AB242016 represent nucleic acid
CC molecules of the invention
XX Sequence 720 BP; 155 A; 245 C; 199 G; 121 T; 0 U; 0 Other;

Query Match: 75.9%; Score 17; DB 10; Length 720;
Best Local Similarity 60.0%; Pred. No. 3.2e+02;
Matches 15; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 3 CGGAUGCGGUCDCDADGDTGDD 27
DB 88 CGGATCGGTCAGCAGCGGTAGA 64

RESULT 9
ID AB241652
XX AB241652 standard; DNA; 765 BP.
AC AB241652;
XX 07-MAR-2003 (first entry)
XX N. gonorrhoeae nucleotide sequence SEQ ID 7893.
XX

KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX Neisseria gonorrhoeae.
XX WO200279243-A2.
XX 10-OCT-2002.
XX 12-FEB-2002; 2002WO-IB002069.
XX 12-FEB-2001; 2001GB-00003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Massignani V, Monaci E;
DR WPI; 2003-058415/05.
DR P-PSDB; ABP80682.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
medicament for treating or preventing N. gonorrhoeae infection.
PS Disclosure; Page 763-764; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records AB237706-AB242016 represent nucleic acid
CC molecules of the invention
XX Sequence 765 BP; 131 A; 208 C; 259 G; 167 T; 0 U; 0 Other;

Query Match: 75.9%; Score 17; DB 10; Length 765;
Best Local Similarity 60.0%; Pred. No. 3.2e+02;
Matches 15; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 3 CGGAUGCGGUCDCDADGDTGDD 27
DB 654 CGGATCGGTCAGCAGCGGTAGA 678

RESULT 10
ACL70886
ID ACL70886 standard; DNA; 1377 BP.
XX ACL70886;
XX 02-JUN-2005 (first entry)
XX M. xanthus gene sequence, seq id 7349.
XX Transgenic plant; DNA replication; gene regulation; gene expression;
KW gene; ds.
XX Myxococcus xanthus.
XX US6833447-B1.
XX 21-DEC-2004.
XX 10-JUL-2001; 2001US-00902540.
XX 10-JUL-2000; 2000US-0217883P.
XX (MONS) MONSANTO TECHNOLOGY LLC.
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX

PT New substantially purified Myxococcus xanthus nucleic acid molecule
 PT encoding a nitrite reductase, useful for determining gene expression,
 PT identifying mutations in a gene of interest, and for constructing
 PT mutations in a gene of interest.

XX Example 2; SEQ ID NO 7349; 25pp; English.

PS The invention relates to a substantially purified nucleic acid molecule
 CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
 CC recombinant DNA construct for expression of a nitrite reductase gene in a
 CC plant cell, and a plant cell comprising the recombinant DNA construct.
 CC The nucleic acid is useful for determining gene expression, identifying
 CC mutations in a gene of interest, and for constructing mutations in a gene
 CC of interest. Sequences given in records for SEQ IDs 1850-9691 represent a
 CC set of about 7842 genes or partial genes from the genome of the bacterium
 CC Myxococcus xanthus. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from USPTO

XX Sequence 1377 BP; 243 A; 453 C; 485 G; 196 T; 0 U; 0 Other;

Query Match 74.1%; Score 16.6; DB 14; Length 1377;
 Best Local Similarity 55.6%; Pred. No. 5.5e+02;
 Matches 15; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUDCDADGDTGDD 27

Db 525 GAAGGATCGGCTTCGCCCGGTGGG 551

RESULT 11

ACL64238/C
 ID ACL64238 standard; DNA; 5515 BP.

XX ACL64238;

DT 02-JUN-2005 (first entry)

DE M. xanthus DNA fragment, seq id 701.

XX Transgenic plant; DNA replication; gene regulation; gene expression; ds.

OS Myxococcus xanthus.

XX US6833447-B1.

XX 21-DEC-2004.

XX 10-JUL-2001; 2001US-00902540.

XX 10-JUL-2000; 2000US-0217883P.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;

XX WPI; 2005-028716/03.

XX New substantially purified Myxococcus xanthus nucleic acid molecule
 PT encoding a nitrite reductase, useful for determining gene expression,
 PT identifying mutations in a gene of interest, and for constructing
 PT mutations in a gene of interest.

PS Example 1; SEQ ID NO 701; 25pp; English.

XX The invention relates to a substantially purified nucleic acid molecule
 CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
 CC recombinant DNA construct for expression of a nitrite reductase gene in a
 CC plant cell, and a plant cell comprising the recombinant DNA construct.
 CC The nucleic acid is useful for determining gene expression, identifying
 CC mutations in a gene of interest, and for constructing mutations in a gene
 CC of interest. Sequences given in records for SEQ IDs 1-1849 represent a
 CC set of 1849 contig and singleton sequences comprising coding sequences,

CC DNA replication elements, promoters and other regulatory elements from
 CC the genome of the bacterium Myxococcus xanthus. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO

XX Sequence 5515 BP; 913 A; 1924 C; 1809 G; 869 T; 0 U; 0 Other;

Query Match 74.1%; Score 16.6; DB 14; Length 5515;
 Best Local Similarity 55.6%; Pred. No. 6.5e+02;
 Matches 15; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUDCDADGDTGDD 27

Db 3521 GAAGGATCGGCTTCGCCCGGTGGG 3495

RESULT 12

ADN46845_05

Continuation (6 of 21) of ADN46845 from base 500001 (Thermococcus kodakaraensis KOD1 gene
 WP Sequence split into 21 fragments LOCUS ADN46845 Accession Adn46845

WP	Fragment Name	Begin	End
WP	ADN46845_00	1	110000
WP	ADN46845_01	100001	210000
WP	ADN46845_02	200001	310000
WP	ADN46845_03	300001	410000
WP	ADN46845_04	400001	510000
WP	ADN46845_05	500001	610000
WP	ADN46845_06	600001	710000
WP	ADN46845_07	700001	810000
WP	ADN46845_08	800001	910000
WP	ADN46845_09	900001	1010000
WP	ADN46845_10	1000001	1110000
WP	ADN46845_11	1100001	1210000
WP	ADN46845_12	1200001	1310000
WP	ADN46845_13	1300001	1410000
WP	ADN46845_14	1400001	1510000
WP	ADN46845_15	1500001	1610000
WP	ADN46845_16	1600001	1710000
WP	ADN46845_17	1700001	1810000
WP	ADN46845_18	1800001	1910000
WP	ADN46845_19	1900001	2010000
WP	ADN46845_20	2000001	2089378

Query Match 74.1%; Score 16.6; DB 12; Length 110000;
 Best Local Similarity 55.6%; Pred. No. 9.4e+02;
 Matches 15; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACGGAUGCGGUCUDCDADGDTGDDT 28

Db 33146 AAGGATCGGCTTCCTGACGATGTT 33172

RESULT 13

ADN47591_15/c

Continuation (16 of 21) of ADN47591 from base 1500001 (Thermococcus kodakaraensis KOD1 gene
 WP Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591

WP	Fragment Name	Begin	End
WP	ADN47591_00	1	110000
WP	ADN47591_01	100001	210000
WP	ADN47591_02	200001	310000
WP	ADN47591_03	300001	410000
WP	ADN47591_04	400001	510000
WP	ADN47591_05	500001	610000
WP	ADN47591_06	600001	710000
WP	ADN47591_07	700001	810000
WP	ADN47591_08	800001	910000
WP	ADN47591_09	900001	1010000
WP	ADN47591_10	1000001	1110000
WP	ADN47591_11	1100001	1210000
WP	ADN47591_12	1200001	1310000
WP	ADN47591_13	1300001	1410000
WP	ADN47591_14	1400001	1510000
WP	ADN47591_15	1500001	1610000

WP ADN47591_16 1600001 1710000
WP ADN47591_17 1700001 1810000
WP ADN47591_18 1800001 1910000
WP ADN47591_19 1900001 2010000
WP ADN47591_20 2000001 2089378

Query Match 74.1%; Score 16.6; DB 12; Length 110000;
Best Local Similarity 55.6%; Pred. No. 9.4e+02;
Matches 15; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACGGAUGCGGUCDCDADGDTGDT 28
DB 56232 AAGGATGGGCTCTCTGACGATTGTT 56206

RESULT 14

ADN46123_05
Continuation (6 of 21) of ADN46123 from base 500001 (Thermococcus kodakaraensis KOD1 Gen
WP Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123

WP	Fragment Name	Begin	End
WP	ADN46123_00	1	110000
WP	ADN46123_01	100001	210000
WP	ADN46123_02	200001	310000
WP	ADN46123_03	300001	410000
WP	ADN46123_04	400001	510000
WP	ADN46123_05	500001	610000
WP	ADN46123_06	600001	710000
WP	ADN46123_07	700001	810000
WP	ADN46123_08	800001	910000
WP	ADN46123_09	900001	1010000
WP	ADN46123_10	1000001	1110000
WP	ADN46123_11	1100001	1210000
WP	ADN46123_12	1200001	1310000
WP	ADN46123_13	1300001	1410000
WP	ADN46123_14	1400001	1510000
WP	ADN46123_15	1500001	1610000
WP	ADN46123_16	1600001	1710000
WP	ADN46123_17	1700001	1810000
WP	ADN46123_18	1800001	1910000
WP	ADN46123_19	1900001	2010000
WP	ADN46123_20	2000001	2089378

Query Match 74.1%; Score 16.6; DB 12; Length 110000;
Best Local Similarity 55.6%; Pred. No. 9.4e+02;
Matches 15; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACGGAUGCGGUCDCDADGDTGDT 28
DB 33146 AAGGATGGGCTCTCTGACGATTGTT 33172

RESULT 15

ADN47209_15/c
Continuation (16 of 21) of ADN47209 from base 1500001 (Thermococcus kodakaraensis KOD 9
WP Sequence split into 21 fragments LOCUS ADN47209 Accession Adn47209

WP	Fragment Name	Begin	End
WP	ADN47209_00	1	110000
WP	ADN47209_01	100001	210000
WP	ADN47209_02	200001	310000
WP	ADN47209_03	300001	410000
WP	ADN47209_04	400001	510000
WP	ADN47209_05	500001	610000
WP	ADN47209_06	600001	710000
WP	ADN47209_07	700001	810000
WP	ADN47209_08	800001	910000
WP	ADN47209_09	900001	1010000
WP	ADN47209_10	1000001	1110000
WP	ADN47209_11	1100001	1210000
WP	ADN47209_12	1200001	1310000
WP	ADN47209_13	1300001	1410000
WP	ADN47209_14	1400001	1510000
WP	ADN47209_15	1500001	1610000
WP	ADN47209_16	1600001	1710000

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2006, 14:45:52 ; Search time 1731.48 Seconds
(without alignments)
756.599 Million cell updates/sec

Title: US-10-824-829-1

Perfect score: 22.4

Sequence: 1 gacgaugcgguccddcdadgdtdgdt 28

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_ges1:*

10: gb_ges2:*

11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	82.1	396	7	CK579577 IST W15 2
2	18.4	82.1	582	6	CF268393 TgEST27J2
3	18.4	82.1	709	3	BM617570 170006871
4	18.2	81.2	621	6	CF445999 EST682344
5	18.2	81.2	789	6	CF447839 EST684184
6	18.2	81.2	888	9	CG128124 NDL.8117
7	18.2	80.4	368	9	AO639502 927P1-4B1
8	18.0	80.4	491	10	CG047300 PUMK57TB
9	18.0	80.4	804	9	BZ684578 PUBIX92TD
10	18.0	80.4	982	11	AL399350 T3 end of
11	17.6	78.6	307	11	CR089064 Forward s
12	17.6	78.6	482	7	CO140182 EST834853
13	17.6	78.6	576	7	CO147122 EST822175
14	17.6	78.6	621	6	CD786254 EST657615
15	17.6	78.6	640	7	CO132625 GR_Eb45J
16	17.6	78.6	663	2	BB644925 BB644925
17	17.6	78.6	668	11	CR918052 Sus scrofa
18	17.6	78.6	689	7	CO148780 EST823833
19	17.6	78.6	701	7	CK185359 EST774674
20	17.6	78.6	701	7	CK185360 EST774675
21	17.6	78.6	706	10	CG409802 tigr-g88-
22	17.6	78.6	727	7	CO108153 GR_Eb003

C	23	17.6	78.6	746	7	CO081366	GR_Ea45G
	24	17.6	78.6	767	7	CO081337	GR_Ea45F
	25	17.6	78.6	774	10	CZ346745	ZMBBF0120
	26	17.6	78.6	816	7	CO108154	GR_Eb003
	27	17.6	78.6	857	7	CO073821	GR_Ea33J
	28	17.6	78.6	866	6	CD795858	EST667219
	29	17.6	78.6	867	9	CZ124323	NDL_6C6_T
	30	17.6	78.6	875	10	CZ211497	AI2A-aad9
C	31	17.6	78.6	878	6	CD795857	EST667218
	32	17.6	78.6	893	9	CC132402	NDL_92H2.
	33	17.6	78.6	903	7	CO087751	GR_Ea06F
	34	17.6	78.6	946	4	AK046315	Mus muscu
C	35	17.4	77.7	510	10	CL758634	CL758634
	36	17.4	77.7	558	8	DN839607	OR_BBa012
	37	17.4	77.7	621	10	CL612241	CL612241
C	38	17.4	77.7	624	5	BO601073	MI-F-H1-a
	39	17.4	77.7	705	10	AG089856	Pan trogl
	40	17.4	77.7	849	10	CL719417	OR_BBa004
C	41	17.4	77.7	890	10	AG861261	Oryza sat
	42	17.2	76.8	455	3	BM174624	Tm_ad_31G
	43	17.2	76.8	506	1	AI726738	BNLHG1644
	44	17.2	76.8	690	6	CB013537	Tm_ad_60H
	45	17.2	76.8	717	10	CZ701989	OC_Ba001

ALIGNMENTS

RESULT 1

CK579577/c
LOCUS CK579577 396 bp mRNA linear EST 16-JAN-2004
DEFINITION IST W15 25461 AD-wrmcDNA library Caenorhabditis elegans cDNA 5', similar to Y10588A.16, mRNA sequence.
ACCESSION CK579577.1 GI:40963245
VERSION EST.
KEYWORDS Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Pelodierinae; Caenorhabditis.
AUTHORS Li,S., Armstrong,C.M., Bertin,N., Ge,H., Milstein,S., Boxem,M., Vidalain,P.O., Han,J.D., Chesneau,A., Hao,T., Goldberg,D.S., Li,N., Martinez,M., Rual,J.F., Lamesch,P., Xu,L., Tewari,M., Wong,S.L., Zhang,L.V., Berriz,G.F., Jacotot,L., Vaglio,P., Reboul,J., Hirozane-Kishikawa,T., Li,Q., Gabel,H.W., Elewa,A., Baumgartner,B., Rose,D.J., Yu,H., Bosak,S., Sequerra,R., Fraser,A., Mango,S.E., Saxton,M.M., Stromer,S., Van Den Heuvel,S., Piano,P., Vandenhaute,J., Sardet,C., Gerstein,M., Doucette-Stamm,L., Gunsalus,K.C., Harper,J.W., Cusick,M.E., Roth,P.P., Hill,D.E. and Vidal,M.
A Map of the Interactome Network of the Metazoan C. elegans
Science (2004) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
For the purpose of protein interaction mapping, we generated a C. elegans cDNA library (AD-wrmcDNA) in which poly(dT)-primed reverse transcribed cDNA are fused to the AD-encoding sequence of the yeast transcription factor GAL4. cDNAs were generated and cloned into the two hybrid vector pPC86 This Interacting Sequence Tag IST_W15_25461 (Y10588A.16) interacts as a prey with the bait C5662.7
PCR Primers
FORWARD: CGCGTTTGAATCACTACAGGG
BACKWARD: GGAGACTTGACCAACTCTGGCG
Insert Length: 396 Std Error: 264.00
Plate: 342 row: 02 column: B
Seq primer: CGCGTTTGAATCACTACAGGG
High quality sequence stop: 395

TITLE

Science (2004) In press

JOURNAL

COMMENT

```

/db_xref="taxon:5811"
/clone="TgBSTzyj24d08.y1"
/dev_stage="Tachyzoite"
/lab_host="ElectroTen Blue cells (Stratagene)"
/clone_lib="TgWAS Tachyzoite cDNA Library"
/note="Vector: pBluescript II SK+; Site1: EcoRI; Site 2: XhoI; The cDNA library was constructed by Keliang Tang, and Robert Cole at Washington University. cDNA was synthesized from Poly(A)+ mRNA using an oligo-d(T) primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA, and products were size-selected on sephacryl S500. The cDNA were directionally cloned into the EcoRI/XhoI prepared pBluescript II SK+ vector, and electroporated into ElectroTen Blue cells (Stratagene). The library may contain a small percentage of host or bacterial contaminants."

```

ORIGIN

	Query Match	82.1%	Score 18.4;	DB 6;	Length 582;
	Best Local Similarity	62.5%;	Pred. No. 4.5e+02;		
	Matches 15;	Conservative 8;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	GACGGAGCGGTCUCDCDADGDT	24		
db	521	GACGGATCGGTCCTTGACGATGTT	544		

RESULT 3	BM617570	LOCUS	709 bp	mRNA	linear	EST 25-FEB-2002
DEFINITION	17000687159695	A.Gam.ad.cDNA.Blood1				
	196004497114411	5'. mRNA sequence.				

ACCESSION BM617570
VERSION BM617570.1 GI:18915792
KEYWORDS EST.

KEYWORDS EST. *Anopheles gambiae* (African malaria mosquito) SOURCE

ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;
Culicidae; Anophelinae; Anopheles.
REFERENCE 1 (bases 1 to 709)
AUTHORS Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L.,
Charlab, R., Collins, P.H., Venter, J.C. and Hoffman, S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.

Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580

Email: HoltraA@celera.com
Plate: NU01004AXJ row: A column: 21
seq primer: M13 Reverse.

FEATURES

source
1. .709

```

/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.

```

chromosome) "

```
/db_xref="taxon:7165"
```

```
/clone="19600449714411"
```

```
/dev_stage="Adult"
```

```

/lab_host="DH10b"

```

```
/clone lib="A.Gam.ad.cDNA.blood1"
/notc=Vector; Site 1: SalI; Site 2: NotI;
/site=Vector; Site 1: SalI; Site 2: NotI;
adult mosquitoes (mixed sex)_frozen on liquid nitro-
hours after human blood feeding. cDNA inserts >500
cloned directionally into pSport 1. Not 1 site is
```

Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org).


```

Query Match      82.1%; Score 18.4; DB 3; Length 709;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GACGAGCGGUCUCCDCCDAGDGTGDT 28
    |||||:||||:||||:||||:||||:
Db 71 GCGGATCGGTCTACGCAACATCGGT 98
    |||||:||||:||||:||||:||||:

RESULT 4
CF445999
LOCUS      EST682344 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACARV13, mRNA sequence.
ACCESSION  CF445999
VERSION     CF445999.1 GI:34468701
KEYWORDS    EST.
SOURCE      Allium cepa (onion)
ORGANISM    Allium cepa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
            Allium.
REFERENCE   1 (bases 1 to 621)
AUTHORS    Havey,M.J., Cheung,F., Van Aken,S., Utterback,T. and Town,C.D.
TITLE      Expressed Sequence Tags from a normalized library of mixed onion
            tissues (Allium cepa)
JOURNAL    Unpublished (2003)
COMMENT    Contact: Havey MJ
            Department of Horticulture
            USDA-ARS and University of Wisconsin
            1575 Linden Drive, Madison, WI 53706, USA
            Tel: 608-262-1830
            Fax: 608-262-4743
            Email: mhavey@facstaff.wisc.edu
            TIGR sequence name ACARV13TR. For more information:
            http://haveylab.hort.wisc.edu
            Seq primer: CAG GAA ACA GCT ATG ACC.
            Location/Qualifiers
                1..621
                /organism="Allium cepa"
                /mol_type="mRNA"
                /cultivar="Red Creole (bulbs), unknown(callus), Ebano &
                Texas Legend (roots)"
                /db_xref="taxon:4679"
                /clone="ACARV13"
                /tissue type="Callus, roots, and young bulbs"
                /clone_lib="normalized cDNA library of onion"
                /note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
                EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
                from callus, roots, and young bulbs were combined to
                synthesize the library. Normalization to enrich for
                low-copy transcripts was performed by proprietary
                techniques of Invitrogen."

FEATURES
    source
    1..621

ORIGIN
Query Match      81.2%; Score 18.2; DB 6; Length 621;
Best Local Similarity 55.6%; Pred. No. 5.7e+02;
Matches 15; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GACGAGCGGUCUCCDCCDAGDGTGDT 27
    |||||:||||:||||:||||:||||:
Db 231 GACGATCGGTCTATCGAGTTTGA 257
    |||||:||||:||||:||||:||||:

RESULT 5
CF447839
LOCUS      EST684184 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACARF24, mRNA sequence.
ACCESSION  CF447839
VERSION     CF447839.1 GI:34470541
KEYWORDS    EST.
SOURCE      Allium cepa (onion)
ORGANISM    Allium cepa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
            Allium.
REFERENCE   1 (bases 1 to 621)
AUTHORS    Havey,M.J., Cheung,F., Van Aken,S., Utterback,T. and Town,C.D.
TITLE      Expressed Sequence Tags from a normalized library of mixed onion
            tissues (Allium cepa)
JOURNAL    Unpublished (2003)
COMMENT    Contact: Havey MJ
            Department of Horticulture
            USDA-ARS and University of Wisconsin
            1575 Linden Drive, Madison, WI 53706, USA
            Tel: 608-262-1830
            Fax: 608-262-4743
            Email: mhavey@facstaff.wisc.edu
            TIGR sequence name ACARV13TR. For more information:
            http://haveylab.hort.wisc.edu
            Seq primer: CAG GAA ACA GCT ATG ACC.
            Location/Qualifiers
                1..621
                /organism="Allium cepa"
                /mol_type="mRNA"
                /cultivar="Red Creole (bulbs), unknown(callus), Ebano &
                Texas Legend (roots)"
                /db_xref="taxon:4679"
                /clone="ACARV13"
                /tissue type="Callus, roots, and young bulbs"
                /clone_lib="normalized cDNA library of onion"
                /note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
                EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
                from callus, roots, and young bulbs were combined to
                synthesize the library. Normalization to enrich for
                low-copy transcripts was performed by proprietary
                techniques of Invitrogen."

ORIGIN
Query Match      81.2%; Score 18.2; DB 6; Length 621;
Best Local Similarity 55.6%; Pred. No. 5.7e+02;
Matches 15; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GACGAGCGGUCUCCDCCDAGDGTGDT 27
    |||||:||||:||||:||||:||||:
Db 231 GACGATCGGTCTATCGAGTTTGA 257
    |||||:||||:||||:||||:||||:

RESULT 5
CF447839
LOCUS      EST684184 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACARF24, mRNA sequence.
ACCESSION  CF447839
VERSION     CF447839.1 GI:34470541
KEYWORDS    EST.
SOURCE      Allium cepa (onion)
ORGANISM    Allium cepa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
            Allium.
REFERENCE   1 (bases 1 to 789)
AUTHORS    Havey,M.J., Cheung,F., Van Aken,S., Utterback,T. and Town,C.D.
TITLE      Expressed Sequence Tags from a normalized library of mixed onion
            tissues (Allium cepa)
JOURNAL    Unpublished (2003)
COMMENT    Contact: Havey MJ
            Department of Horticulture
            USDA-ARS and University of Wisconsin
            1575 Linden Drive, Madison, WI 53706, USA
            Tel: 608-262-1830
            Fax: 608-262-4743
            Email: mhavey@facstaff.wisc.edu
            TIGR sequence name ACARF24TR. For more information:
            http://haveylab.hort.wisc.edu
            Seq primer: CAG GAA ACA GCT ATG ACC.
            Location/Qualifiers
                1..789
                /organism="Allium cepa"
                /mol_type="mRNA"
                /cultivar="Red Creole (bulbs), unknown(callus), Ebano &
                Texas Legend (roots)"
                /db_xref="taxon:4679"
                /clone="ACARF24"
                /tissue type="Callus, roots, and young bulbs"
                /clone_lib="normalized cDNA library of onion"
                /note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
                EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
                from callus, roots, and young bulbs were combined to
                synthesize the library. Normalization to enrich for
                low-copy transcripts was performed by proprietary
                techniques of Invitrogen."

ORIGIN
Query Match      81.2%; Score 18.2; DB 6; Length 789;
Best Local Similarity 55.6%; Pred. No. 5.8e+02;
Matches 15; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GACGAGCGGUCUCCDCCDAGDGTGDT 27
    |||||:||||:||||:||||:||||:
Db 666 GACGATCGGTCTATCGAGTTTGA 692
    |||||:||||:||||:||||:||||:

RESULT 6
CC128124/c
LOCUS      NDL.8117.T7 Notre Dame Liverpool Aedes aegypti genomic clone
DEFINITION NDL.8117, genomic survey sequence.
ACCESSION  CC128124
VERSION     CC128124.1 GI:29997179
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
ORGANISM    Aedes aegypti
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Culicidae; Culicini; Culicini; Aedes; Stegomyia.
REFERENCE   1 (bases 1 to 888)
AUTHORS    Loftus,B., Shetty,J., Knudson,D. and Severson,D.
TITLE      BAC end sequencing of Aedes aegypti
JOURNAL    Unpublished (2003)
COMMENT    Other GSSs: NDL.8117.SP6
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: entaetigr.org
            Library was provided by David Severson
            Seq primer: T7

```

```

FEATURES
  source
    Class: BAC ends.
    Location/Qualifiers
      1..888
        /organism="Aedes aegypti"
        /mol_type="genomic DNA"
        /strain="Liverpool"
        /db_xref="taxon:7159"
        /clone="NDL.8117"
        /note="Vector: pCBAC1; Site 1: Hind III; The library was
        prepared from whole body tissue of newly hatched 1L larvae
        by David Severson at the University of Notre Dame and
        Hongbin Zhang"

ORIGIN
  Query Match      81.2%; Score 18.2; DB 9; Length 888;
  Best Local Similarity 55.6%; Pred. No. 5.9e+02;
  Matches 15; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACAGGAGCGGUCUCCDADGDTGDT 28
DB 286 ACAGATCCGATCTTCGATGATGTT 260

RESULT 7
A0639502
LOCUS
DEFINITION
  A0639502
  927P1-4B10.TP 927P1 Trypanosoma brucei genomic clone 927P1-4B10,
  genomic survey sequence.
ACCESSION
  A0639502
  A0639502.1 GI:5116212
KEYWORDS
  GSS.
SOURCE
  Trypanosoma brucei
  ORGANISM
    Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
    Trypanosoma.
REFERENCE
  1 (bases 1 to 368)
  El-Sayed, N., Zhao, S., Zhao, H., Gull, S., Suh, E., Malek, J., Fujii, C.,
  Gerrard, C., Leech, V., de Jong, P., Ullu, E., Melville, S.,
  Donelson, J., Fraser, C. and Adams, M.
  Determination of clone end sequences from Trypanosoma brucei TREU
  927/4 P1 library
  Unpublished (1999)
  Other GSSs: 927P1-4B10.TV
  Contact: Najib M. El-Sayed
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: nelsayed@tigr.org
  For clone/filter availability, please contact Sara Melville
  (sm160@mol.bio.cam.ac.uk). P1 end sequences search page:
  http://www.tigr.org/tdb/mdb/tbdb/.
  Seq primer: SP6.
  Class: P1 ends.
  Location/Qualifiers
    1..368
      /organism="Trypanosoma brucei"
      /mol_type="genomic DNA"
      /strain="TREU927/4"
      /db_xref="taxon:5691"
      /clone="927P1-4B10"
      /clone_lib="927P1"
      /note="Vector: PAD108acBII; Site 1: Bam HI; Constructed by
      Sara Melville, University of Cambridge, UK and Nancy
      Shepherd, Dupont Merck, Wilmington, DE. Genomic DNA was
      isolated from Trypanosoma brucei (stock TREU927/4) and
      partially digested with Sau 3AI. DNA fragments were cloned
      into the Bam HI site of PAD108acBII vector (Genbank
      accession U09128). The average insert size is 65 Kb.
      Coverage: approx 4.4 X the haploid non-minichromosomal
      genome"

FEATURES
  source
    Class: BAC ends.
    Location/Qualifiers
      1..368
        /organism="Trypanosoma brucei"
        /mol_type="genomic DNA"
        /strain="TREU927/4"
        /db_xref="taxon:5691"
        /clone="927P1-4B10"
        /clone_lib="927P1"
        /note="Vector: PAD108acBII; Site 1: Bam HI; Constructed by
        Sara Melville, University of Cambridge, UK and Nancy
        Shepherd, Dupont Merck, Wilmington, DE. Genomic DNA was
        isolated from Trypanosoma brucei (stock TREU927/4) and
        partially digested with Sau 3AI. DNA fragments were cloned
        into the Bam HI site of PAD108acBII vector (Genbank
        accession U09128). The average insert size is 65 Kb.
        Coverage: approx 4.4 X the haploid non-minichromosomal
        genome"

```

ORIGIN

```

Query Match      80.4%; Score 18; DB 9; Length 368;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 16; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUCCDADGDTG 26
DB 276 GACGGAAGCGTCTTCGCAATGTTG 301

```

RESULT 8

```

CG047300
LOCUS
DEFINITION
  CG047300
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
  Location/Qualifiers
    1..491
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBTA0616L13"
      /clone_lib="ZM 0.6 1.0 KB"
      /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
      Cot selected genomic DNA library"

```

ACCESSION

```

CG047300
CG047300.1 GI:33919480

```

KEYWORDS

```

GSS.

```

SOURCE

```

Zea mays

```

ORGANISM

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

```

REFERENCE

```

1 (bases 1 to 491)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUINK67TD
Contact: Cathy Whitelaw
TIGR

```

TITLE

```

JOURNAL
COMMENT

```

FEATURES

```

source
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
  Location/Qualifiers
    1..491
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBTA0616L13"
      /clone_lib="ZM 0.6 1.0 KB"
      /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
      Cot selected genomic DNA library"

```

ORIGIN

```

Query Match      80.4%; Score 18; DB 10; Length 491;
Best Local Similarity 61.5%; Pred. No. 7e+02;
Matches 16; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUCCDADGDTG 26
DB 345 GACGGATACGGTCTCAAGGATGG 370

```

RESULT 9

```

BZ684578/c
LOCUS
DEFINITION
  BZ684578
  PUBIX92TD ZM_0.6 1.0 KB Zea mays genomic clone ZMMBTA065P16,
  genomic survey sequence.
ACCESSION
  BZ684578
  BZ684578.1 GI:28243043
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.

```


JOURNAL COMMENT

FEEMS Microbiol. Lett. (2004) In press

Contact: Yu J

Food and Feed Safety Research Unit

USDA/ARS, Southern Regional Research Center

1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA

Tel: 504 286 4405

Fax: 504 286 4419

Email: jiuju@rrrc.ars.usda.gov

Contact Dr. Yu at USDA/ARS SRRRC (jiuju@rrrc.ars.usda.gov) for clone information

PCR Primers

FORWARD: M13F

BACKWARD: M13R

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

1. 576

/organism="Aspergillus flavus"

/mol_type="mRNA"

/strain="NRRL 3357"

/db_xref="taxon:5059"

/clone="NAGCS30"

/sex="asexual mycelia"

/cell_type="mycelia"

/dev_stage="developmental stages from 18 to 96 hours"

/lab_host="E. coli DH10B T1 resistant cells"

/clone_lib="Aspergillus flavus Normalized cDNA Expression Library"

/note="Vector: pBlueScript (SK+) (Stratagene), antibiotic selection marker: Carbenicillin; Site1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

ORIGIN

Query Match 78.6%; Score 17.6; DB 7; Length 576;

Best Local Similarity 53.6%; Pred. No. 1.1e+03;

Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAGCGAUGCGGUCUCCDADGDTGDT 28

|||||:||||: : : |:|:|:

Db 446 GACGGATGCGGTATTATGATGTAGAT 473

RESULT 14

CD786254/c

LOCUS

DEFINITION

CD786254 621 bp mRNA linear EST 01-JUL-2004

EST657615 RAA Rhipicephalus appendiculatus CDNA clone RAABP46 3'

end, mRNA sequence.

CD786254

CD786254.1 GI:49541927

EST.

Rhipicephalus appendiculatus

Rhipicephalus appendiculatus

Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Rhipicephalus.

1 (bases 1 to 621)

Nene,V. Lee,D., Kang'a,S., Skilton,R., Shah,T., de Villiers,E., Mwaura,S., Taylor,D., Quackenbush,J. and Bishop,R. Genes transcribed in the salivary glands of female Rhipicephalus appendiculatus ticks infected with Theileria parva Insect Biochem. Mol. Biol. 34 (10), 1117-1128 (2004)

15475305

Contact: Vishvanath Nene

Parasite Genomics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-610-5968

Fax: 301-838-0208

Email: nene@tigr.org

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBLISHED

COMMENT

FEATURES

source

Seq primer: mixed oligo dT.
Location/Qualifiers
1..621
/organism="Rhhipicephalus appendiculatus"
/mol_type="mRNA"
/strain="Muguga"
/db_xref="taxon:34631"
/clone="RAABP46"
/dev_stage="Adult"
/lab_host="B. coli strain DH10B-Tona"
/clone_lib="RAA"
/note="Organ: Salivary glands; Vector: pCMVSPORT6.0.ccdB;
Salivary Glands were dissected on day four after
initiation of feeding. Total RNA was prepared using acid
guanidium thiocyanate-phenol-chloroform extraction. The
cDNA library was custom prepared by Invitrogen
Corporation. Briefly, first strand cDNA was primed using
oligo(dT) containing a NotI site. Size fractionated double
stranded cDNA was ligated to EcoRV-NotI cleaved vector and
electroporated into E.coli. Library RAA was made from
uninfected ticks."

ORIGIN

Query Match 78.6%; Score 17.6; DB 6; Length 621;
Best Local Similarity 53.6%; Pred. No. 1.1e+03;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
QY 1 'GACGGAUGCGGUCUCDCDADGDTGDT 28
|||||:|||||:|||||:|||||:|||||:
Db 341 GACAGATGGGTCTACACAGTGTCTGTT 314

RESULT 15

COI32625

LOCUS

COI32625 GR_Eb45J09.r GR_Eb Gossypium raimondii cDNA clone GR_Eb45J09 3',
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Gossypium raimondii
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE

'AUTHORS

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
Wing, R.A.

TITLE

JOURNAL

COMMENT

Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 45 row: J column: 09.

FEATURES

source

Location/Qualifiers
1..640
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb45J09"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_EB"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 78.6%; Score 17.6; DB 7; Length 640;
Best Local Similarity 53.6%; Pred. No. 1.1e+03;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
QY 1 GACGGAUGCGGUCUCDCDADGDTGDT 28
|||||:|||||:|||||:|||||:|||||:
Db 507 GACGGATGGGTCTTCTCAATTGATGAT 534

Search completed: February 19, 2006, 16:24:05
Job time : 1734.48 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic search, using sw model

Run on: February 18, 2006, 15:01:08 ; Search time 97.5484 Seconds
(without alignments)
510.226 Million cell updates/sec

Title: US-10-824-829-1

Perfect score: 22.4

Sequence: 1 gacggaugcgguccdcadgdtdgdt 28

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCRB COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.6	78.6	1149	3	US-09-533-559-703
2	16.6	74.1	1377	3	US-09-502-540-7349
3	16.6	74.1	5515	3	US-09-902-540-701
4	16.2	72.3	1629	3	US-09-604-231-19
5	16.2	72.3	1827	3	US-09-604-231-17
6	16.2	72.3	2005	3	US-09-950-788B-1
7	16.2	72.3	2005	3	US-09-950-788B-3
8	16.2	72.3	2298	3	US-09-950-788B-6
9	16	71.4	601	3	US-09-949-016-137825
10	16	71.4	601	3	US-09-949-016-137826
11	16	71.4	601	3	US-09-949-016-137827
12	16	71.4	1197	3	US-08-935-263-11
13	16	71.4	1197	3	US-09-594-185-11
14	16	71.4	1197	3	US-10-033-078-11
15	16	71.4	1197	3	US-10-763-933-11
16	16	71.4	2184	3	US-09-381-849-1
17	16	71.4	135171	3	US-09-949-016-15617
18	15.8	70.5	601	3	US-09-949-016-88305
19	15.8	70.5	1710	3	US-08-969-046-3
20	15.8	70.5	1710	3	US-09-530-663B-3
21	15.8	70.5	3985	3	US-09-949-016-1442
22	15.8	70.5	88490	3	US-09-949-016-12758
23	15.8	70.5	88736	3	US-09-949-016-14222
24	15.8	70.5	4403765	3	US-09-103-840A-2

Sequence 1, Appli
Sequence 7561, Ap
Sequence 7923, Ap
Sequence 10428, A
Sequence 35, Appl
Sequence 1636, Ap
Sequence 1672, Ap
Sequence 12956, A
Sequence 14184, A
Sequence 8799, Ap
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 11, Appl
Sequence 6, Appli
Sequence 8992, Ap
Sequence 9100, Ap
Sequence 935, App

c-25 15.8 70.5 4411529 3 US-09-103-840A-1
26 15.6 69.6 723 3 US-09-252-991A-7561
c 27 15.6 69.6 1008 3 US-09-252-991A-7923
c 28 15.6 69.6 1171 3 US-09-270-767-10428
c 29 15.6 69.6 1334 3 US-09-668-097A-35
c 30 15.6 69.6 1686 3 US-09-489-039A-1636
c 31 15.6 69.6 1725 3 US-09-489-039A-1672
32 15.6 69.6 18122 3 US-09-949-016-12956
33 15.6 69.6 183112 3 US-09-949-016-14184
34 15.4 68.7 1272 3 US-09-902-540-8799
35 15.4 68.7 1320 2 US-08-641-038A-1
36 15.4 68.7 1320 2 US-09-059-178-1
37 15.4 68.7 1369 2 US-08-642-541-1
c 38 15.4 68.7 1369 2 US-08-642-541-3
39 15.4 68.7 1369 3 US-09-260-889-1
c 40 15.4 68.7 1369 3 US-09-260-889-3
41 15.4 68.7 1369 3 US-09-347-878-11
42 15.4 68.7 1369 3 US-09-479-275-6
c 43 15.4 68.7 1494 3 US-09-252-991A-8992
44 15.4 68.7 17374 3 US-09-252-991A-9100
45 15.4 68.7 10096 3 US-09-902-540-935

ALIGNMENTS

RESULT 1
US-09-533-559-703/c
; Sequence 703, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 703
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1149)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-703

Query Match 78.6%; Score 17.6; DB 3; Length 1149;
Best Local Similarity 53.6%; Pred. No. 44;
Matches 15; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGGAUGCGGUCCDCADGDGDT 28
Db 664 GATGATGCGCTCTCTCGAGCATTTGGT 637

RESULT 2
US-09-902-540-7349
; Sequence 7349, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

```
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 7349
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7349

Query Match      74.1%; Score 16.6; DB 3; Length 1377;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 15; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACGGAUGCGGUCUDCCDADGDTGDD 27
Db 525 GAAGGATCGGTCTTCGCCCGCTGGG 551

RESULT 3
US-09-902-540-701/C
; Sequence 701, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 701
; LENGTH: 5515
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-701

Query Match      74.1%; Score 16.6; DB 3; Length 5515;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 15; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACGGAUGCGGUCUDCCDADGDTGDD 27
Db 3521 GAAGGATCGGTCTTCGCCCGCTGGG 3495

RESULT 4
US-09-604-231-19
; Sequence 19, Application US/09604231
; Patent No. 6884614
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: PHOSPHOENOLPYRUVATE:SUGAR PHOSPHOTRANSFERASE
; TITLE OF INVENTION: SYSTEM PROTEINS
; FILE REFERENCE: BGI-122CP
; CURRENT APPLICATION NUMBER: US/09/604,231
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 19
; LENGTH: 1629
; TYPE: DNA
```

```
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(1606)
; OTHER INFORMATION: PRXA01244
US-09-604-231-19

Query Match      72.3%; Score 16.2; DB 3; Length 1629;
Best Local Similarity 52.0%; Pred. No. 2.3e+02;
Matches 13; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGAUGCGGUCUDCCDADGDTGDDT 28
Db 847 GGCTGCGGTCTACTCAAAAGGTGCTT 871

RESULT 5
US-09-604-231-17
; Sequence 17, Application US/09604231
; Patent No. 6884614
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: PHOSPHOENOLPYRUVATE:SUGAR PHOSPHOTRANSFERASE
; TITLE OF INVENTION: SYSTEM PROTEINS
; FILE REFERENCE: BGI-122CP
; CURRENT APPLICATION NUMBER: US/09/604,231
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 17
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1804)
; OTHER INFORMATION: RXN01244
US-09-604-231-17

Query Match      72.3%; Score 16.2; DB 3; Length 1827;
Best Local Similarity 52.0%; Pred. No. 2.4e+02;
Matches 13; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGAUGCGGUCUDCCDADGDTGDDT 28
Db 1045 GGCTGCGGTCTACTCAAAAGGTGCTT 1069

RESULT 6
US-09-950-788B-1
; Sequence 1, Application US/09950788B
; Patent No. 6880187
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: HANS, Stephan
; APPLICANT: SCHISCHKA, Natalie
; APPLICANT: PFERLE, Walter
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE PTISI PROTEIN
; FILE REFERENCE: 213544USO
; CURRENT APPLICATION NUMBER: US/09/950,788B
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: DE10045496.8
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 2005
; TYPE: DNA
```


ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (154)..(1857)
OTHER INFORMATION: ptsI gene
US-09-950-788B-1

Query Match 72.3%; Score 16.2; DB 3; Length 2005;
Best Local Similarity 52.0%; Pred. No. 2.4e+02;
Matches 13; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGAUGCGGUCUCCDADGDTGDT 28
Db 1098 GGCTGGGTCTACTCAAAGGTGCTT 1122

RESULT 7
US-09-950-788B-3

; Sequence 3, Application US/09950788B
; Patent No. 6680187
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: HANS, Stephan
; APPLICANT: SCHISCHKA, Natalie
; APPLICANT: PFEFFERLE, Walter
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE PTSI PROTEIN
; FILE REFERENCE: 213544USO
; CURRENT APPLICATION NUMBER: US/09/950,788B
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: DE10045496.8
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2005
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)..(1857)
; OTHER INFORMATION: Allele ptsI-1547
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (520)
; OTHER INFORMATION: A-G transition
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (562)
; OTHER INFORMATION: C-T transition
US-09-950-788B-3

Query Match 72.3%; Score 16.2; DB 3; Length 2005;
Best Local Similarity 52.0%; Pred. No. 2.4e+02;
Matches 13; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGAUGCGGUCUCCDADGDTGDT 28
Db 1098 GGCTGGGTCTACTCAAAGGTGCTT 1122

RESULT 8
US-09-950-788B-6

; Sequence 6, Application US/09950788B
; Patent No. 6680187
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: HANS, Stephan
; APPLICANT: SCHISCHKA, Natalie
; APPLICANT: PFEFFERLE, Walter
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE PTSI PROTEIN
; FILE REFERENCE: 213544USO
; CURRENT APPLICATION NUMBER: US/09/950,788B
; CURRENT FILING DATE: 2001-09-13

; PRIOR APPLICATION NUMBER: DE10045496.8
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (447)..(2150)
; OTHER INFORMATION: ptsI gene
US-09-950-788B-6

Query Match 72.3%; Score 16.2; DB 3; Length 2298;
Best Local Similarity 52.0%; Pred. No. 2.5e+02;
Matches 13; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGAUGCGGUCUCCDADGDTGDT 28
Db 1391 GGCTGGGTCTACTCAAAGGTGCTT 1415

RESULT 9

US-09-949-016-137825
; Sequence 137825, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137825
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-137825

Query Match 71.4%; Score 16; DB 3; Length 601;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 14; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACGGAUGCGGUCUCCDADGDTGDT 28
Db 362 GATGGTGGGTCTTAAGTATGAGTT 389

RESULT 10

US-09-949-016-137826
; Sequence 137826, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; CURRENT APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 137826
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-137826

Query Match 71.4%; Score 16; DB 3; Length 601;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 14; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUDCCDADGDTGDT 28
 DB 326 GATGGTGGCGTCTTAACATGAGTT 353

RESULT 11
 US-09-949-016-137827
 ; Sequence 137827 Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 137827
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-137827

Query Match 71.4%; Score 16; DB 3; Length 601;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 14; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUDCCDADGDTGDT 28
 DB 58 GATGGTGGCGTCTTAACATGAGTT 85

RESULT 12
 US-08-935-263-11
 ; Sequence 11, Application US/08935263A
 ; Patent No. 6117669
 ; GENERAL INFORMATION:
 ; APPLICANT: Furuichi, Yasuhiro
 ; APPLICANT: Hoshino, Tatsuo
 ; APPLICANT: Kimura, Hitoshi
 ; APPLICANT: Kiyasu, Tatsuya
 ; APPLICANT: Nagahashi, Yoshie
 ; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
 ; FILE REFERENCE: Biotin Genes
 ; CURRENT APPLICATION NUMBER: US/08/935,263A
 ; CURRENT FILING DATE: 1997-09-22
 ; EARLIER APPLICATION NUMBER: EP 96115540.5
 ; PRIOR FILING DATE: 1996-09-27
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 1197
 ; TYPE: DNA

; ORGANISM: Kurthia sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1194)
 US-08-935-263-11

Query Match 71.4%; Score 16; DB 3; Length 1197;
 Best Local Similarity 53.6%; Pred. No. 2.8e+02;
 Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUDCCDADGDTGDT 28
 DB 543 GACGATACGGTCTTCAGCATGGACGGT 570

RESULT 13
 US-09-594-185-11
 ; Sequence 11, Application US/09594185
 ; Patent No. 6363388
 ; GENERAL INFORMATION:
 ; APPLICANT: Furuichi, Yasuhiro
 ; APPLICANT: Hoshino, Tatsuo
 ; APPLICANT: Kimura, Hitoshi
 ; APPLICANT: Kiyasu, Tatsuya
 ; APPLICANT: Nagahashi, Yoshie
 ; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
 ; FILE REFERENCE: Biotin Genes
 ; CURRENT APPLICATION NUMBER: US/09/594,185
 ; CURRENT FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: 08/935,263
 ; PRIOR FILING DATE: 1997-09-22
 ; PRIOR APPLICATION NUMBER: EP 96115540.5
 ; PRIOR FILING DATE: 1996-09-27
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 1197
 ; TYPE: DNA
 ; ORGANISM: Kurthia sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1194)
 US-09-594-185-11

Query Match 71.4%; Score 16; DB 3; Length 1197;
 Best Local Similarity 53.6%; Pred. No. 2.8e+02;
 Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACGGAUGCGGUCUDCCDADGDTGDT 28
 DB 543 GACGATACGGTCTTCAGCATGGACGGT 570

RESULT 14
 US-10-033-078-11
 ; Sequence 11, Application US/10033078
 ; Patent No. 6723544
 ; GENERAL INFORMATION:
 ; APPLICANT: Furuichi, Yasuhiro
 ; APPLICANT: Hoshino, Tatsuo
 ; APPLICANT: Kimura, Hitoshi
 ; APPLICANT: Kiyasu, Tatsuya
 ; APPLICANT: Nagahashi, Yoshie
 ; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
 ; FILE REFERENCE: Biotin Genes
 ; CURRENT APPLICATION NUMBER: US/10/033,078
 ; CURRENT FILING DATE: 2001-12-27
 ; PRIOR APPLICATION NUMBER: 09/594,185
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: EP 96115540.5
 ; PRIOR FILING DATE: 1996-09-27
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 11
LENGTH: 1197
TYPE: DNA
ORGANISM: Kurthia sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1194)
US-10-033-078-11

Query Match 71.4%; Score 16; DB 3; Length 1197;
Best Local Similarity 53.6%; Pred. No. 2.8e+02;
Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GACGGAUGCGGUCDCDADGDTGDT 28
Db 543 GACGATACGGTCTTCAGCATGACGGT 570

RESULT 15
US-10-763-933-11
Sequence 11; Application US/10763933
Patent No. 6955906
GENERAL INFORMATION:
APPLICANT: Furuchi, Yasuhiro
APPLICANT: Hoshino, Tatsuo
APPLICANT: Kimura, Hitoshi
APPLICANT: Kiyasu, Tatsuya
APPLICANT: Nagahashi, Yoshie
TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
FILE REFERENCE: Biotin Genes
CURRENT APPLICATION NUMBER: US/10/763,933
CURRENT FILING DATE: 2004-01-23
PRIOR APPLICATION NUMBER: US/10/033,078
PRIOR FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 09/594,185
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: EP 96115540.5
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 1197
TYPE: DNA
ORGANISM: Kurthia sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1194)
US-10-763-933-11

Query Match 71.4%; Score 16; DB 3; Length 1197;
Best Local Similarity 53.6%; Pred. No. 2.8e+02;
Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GACGGAUGCGGUCDCDADGDTGDT 28
Db 543 GACGATACGGTCTTCAGCATGACGGT 570

Search completed: February 18, 2006, 15:20:39
Job time : 102.548 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2006, 14:38:01 ; Search time 1132.42 Seconds
(without alignments)
1706.680 Million cell updates/sec

Title: US-10-824-829-2

Perfect score: 28.4
Sequence: 1 cguauucgacguacuacutdcdadgdcddt 34

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 5893141 seqs, 2842172553 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sv.*

12: gb_un.*

13: gb_vl.*

14: gb_mtg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28.4	100.0	34	6	CQ898517
2	20.2	71.1	1143	6	E50427 Process for
3	20.2	71.1	2775	1	DQ086153 Bacillus
4	20.2	71.1	2989	1	AB016245 Bacillus
5	20.2	71.1	3045	6	E50424 Process for
6	20.2	71.1	3738	1	AB046355 Bacillus
7	20.2	71.1	4315	1	AB039950 Bacillus
8	20.2	71.1	8189	1	BZ92954 B.subtilis
9	20.2	71.1	200690	1	BZ9122 Bacillus su
10	19.8	69.7	187516	8	AP002813 Homo sapi
11	19.8	69.7	191589	8	AP002992 Homo sapi
12	19.6	69.0	999	6	CQ586245 Sequence
13	19.6	69.0	1394	2	AY075253 Drosophila
14	19.6	69.0	2742	6	CQ609320 Sequence
15	19.6	69.0	2999	6	CQ586244 Sequence
16	19.6	69.0	73617	14	AC018226
17	19.6	69.0	182080	2	AC012373 Drosophila
18	19.6	69.0	188489	2	AC012098 Drosophila

C	19	19.6	69.0	244856	14	AC094247	AC094247 Rattus no
	20	19.6	69.0	258052	14	AC114365	AC114365 Rattus no
	21	19.6	69.0	315109	2	AE003509	AE003509 Drosophila
	22	19.4	68.3	1873	15	AK067227	AK067227 Oryza sat
	23	19.4	68.3	43982	15	AC007289	AC007289 Arabidops
	24	19.4	68.3	78600	15	AP002487	AP002487 Oryza sat
	25	19.4	68.3	85165	15	NC99H12	AL451018 Neurospor
	26	19.4	68.3	110000	14	CR954207_2	Continuation (3 of
	27	19.4	68.3	110000	15	AP008207_011	Continuation (12 o
	28	19.4	68.3	141079	15	AP002868	AP002868 Oryza sat
	29	19.4	68.3	149727	14	AC073808	AL513466 Neurospor
	30	19.4	68.3	157180	15	NCB18D24	AL513466 Mus muscu
	31	19.4	68.3	189055	14	AC073716	AC073716 Mus muscu
	32	19.4	68.3	214343	14	AC155163	AC155163 Mus muscu
	33	19.4	68.3	219630	9	AC161765	AC161765 Mus muscu
	34	19.2	67.6	1366	2	AF432347	AF432347 Macrobac
	35	19.2	67.6	110000	14	CT005269_10	Continuation (11 o
	36	19.2	67.6	110000	14	LMF1CHR32_02	Continuation (3 of
	37	19	66.9	110000	15	AP008211_005	Continuation (6 of
	38	19	66.9	176678	15	AC079356	AC079356 Oryza sat
	39	18.8	66.2	91172	8	AC008553	AC008553 Homo sapi
	40	18.8	66.2	110000	1	CP000091_09	Continuation (10 o
	41	18.8	66.2	110000	14	TANN2_14	Continuation (15 o
	42	18.8	66.2	110000	15	CR382137_14	Continuation (15 o
	43	18.8	66.2	149102	8	AC108095	AC108095 Homo sapi
	44	18.8	66.2	189656	14	AC113421	AC113421 Homo sapi
	45	18.8	66.2	234676	14	AC127042	AC127042 Rattus no

ALIGNMENTS

RESULT 1	CQ898517	Sequence 2 from Patent WO2004092418.	34 bp	DNA	linear	PAT 08-NOV-2004
LOCUS	CQ898517	Sequence 2 from Patent WO2004092418.	34 bp	DNA		
DEFINITION	CQ898517	Sequence 2 from Patent WO2004092418.	34 bp	DNA		
ACCESSION	CQ898517	Sequence 2 from Patent WO2004092418.	34 bp	DNA		
VERSION	CQ898517.1	GI:55582826				
KEYWORDS		synthetic construct				
SOURCE		synthetic construct				
ORGANISM		other sequences; artificial sequences..				
REFERENCE		1				
AUTHORS		Kurn, N. and Wang, S.				
TITLE		Global amplification using a randomly primed composite primer				
JOURNAL		Patent: WO 2004092418-A 2 28-OCT-2004;				
FEATURES		NUGEN Technologies, Inc. (US)				
source		Location/Qualifiers				
		1..34				
		/organism="synthetic construct"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:32630"				
		/note="Primer"				
		1..21				
		/note="Ribonucleotides"				
		22..34				
		/note="Deoxyribonucleotides"				

Query Match	100.0%	Score 28.4;	DB 6;	Length 34;
Best Local Similarity	82.4%	Pred. No. 0.011;		
Matches 28;	Conservative	6;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy	1	CGUAUUCGACGACGACUCUCU	CDTDCDADGDCDDT	34
Db	1	CGTATTTCGACGACGACGAC	CTCTCTDTCADGDCDDT	34

RESULT 2	E50427	1143 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	E50427	1143 bp	DNA	linear	
DEFINITION	E50427	Process for producing poly-gamma-glutamic acid.			
ACCESSION	E50427				

```

VERSION      E50427.1 GI:18629415
KEYWORDS     JP 2001017182-A/4.
SOURCE       Bacillus subtilis
ORGANISM     Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE    1 (bases 1 to 1143)
AUTHORS     Ashiuchi,M., Misono,H. and Soda,K.
TITLE       Process for producing poly-gamma-glutamic acid
JOURNAL     Patent: JP 2001017182-A 4 23-JAN-2001;
            NAGASE & CO LTD
COMMENT      OS Bacillus subtilis (hay bacillus) IFO 3336
            PN JP 2001017182-A/4
            PD 23-JAN-2001
            PF 09-JUL-1999 JP 1999196335
            PR
            PI MAKOTO ASHIUCHI, HARUO MISONO, KENJI SODA
            PC C12N15/09, C08G69/36, C12N1/15, C12N1/19, C12N1/21, C12N5/10, PC
            C12N9/00, C12N9/90,
            PC C12P13/14, C12N15/00, C12N5/00
            CC
            FH Key Location/Qualifiers
            FT source 1..1143
            FT /organism="Bacillus subtilis"
            FT IFO 3336".
FEATURES     source
            Location/Qualifiers
            1..1143
            /organism="Bacillus subtilis"
            /mol_type="genomic DNA"
            /db_xref="taxon:1423"
ORIGIN
Query Match 71.1%; Score 20.2; DB 6; Length 1143;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGUATTCGACGAGCTACTCT 178
DB 158 CGTATTCGACGAGCTACTCT 178
RESULT 3
DQ086153 2775 bp DNA linear BCT 27-JUN-2005
DEFINITION Bacillus subtilis strain ZJU-7 PgsB (pgsB), PgsC (pgsC), and PgsA
            (pgsA) genes, complete cds.
ACCESSION DQ086153
VERSION DQ086153.1 GI:68138277
KEYWORDS Bacillus subtilis
SOURCE Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ORGANISM
REFERENCE 1 (bases 1 to 2775)
AUTHORS Shi, F., Xu, Z. and Cen, P.
TITLE Efficient production of poly(glutamic acid) by a new strain
JOURNAL Bacillus subtilis ZJU-7
REFERENCE 2 (bases 1 to 2775)
AUTHORS Shi, F., Xu, Z. and Cen, P.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2005) Institute of Bioengineering, Department of
            Chemical Engineering and Bioengineering, Zhejiang University, No.38,
            Zheda Road, Hangzhou, Zhejiang 310027, P.R. China
FEATURES     source
            Location/Qualifiers
            1..2775
            /organism="Bacillus subtilis"
            /mol_type="genomic DNA"
            /strain="ZJU-7"
            /db_xref="taxon:1423"
            1..1182
            /gene="pgsB"
            1..1182
            /gene="pgsA"
            /function="Biosynthesis of poly(glutamic acid)"
gene
CDS

```

```

/codon_start=1
/transl_table=11
/product="PgsB"
/protein_id="AA085940.1"
/db_xref="GI:68138278"
/translation="MWLLIACAVILVIGLEKRRHOKNIDALPVRVNINGIRGKSTV
RLTTGILIRAGYTKGTGTGDARMIYWDTPPEKPIKRKPGQNGI GQEKVMEETVE
RGANAYSECMANPDYQIIFQELLOANIGTVNVLEDHMDVNGPTLDETAFAFTAT
IPYNGHLVITDSEYTFEKKAKERNKVIADNSKITDEYLRKEFYVVFEDNASLAL
GVAQALGIDBETAFKMLNAPPDPAARLPLISFSPGPFVNGFAANDASSTLNWK
RVKEIGYPTDDPIIIMNCRADRVDRTOQFANDPLPYIEASELILIGITTEPIVKAYEE
GKIPADKLDHLEVYKSTDEIMELLKRMHNRVYGVGNHGAEEPLIEKIHVKVQLV
S"
1183..1632
/gene="pgsA"
1183..1632
/gene="pgsC"
/function="Biosynthesis of poly(glutamic acid)"
/codon_start=1
/transl_table=11
/product="PgsA"
/protein_id="AA085941.1"
/db_xref="GI:68138279"
/translation="MFQSDLYIALILGVLLSLIFAEKTI GIVPAGLVVPGYLGVPNQF
VFILLVLSLTVYIVKGLSKPMILYGRKKFAAMLTIGVLKIAFDPLYPVPEFI
AEFRGIGIYVPLIANTIQKGLTITFGSTLLSGATPAIMFVYLI"
1633..2775
/gene="pgsA"
1633..2775
/gene="pgsA"
/function="Biosynthesis of poly(glutamic acid)"
/codon_start=1
/transl_table=11
/product="PgsA"
/protein_id="AA085942.1"
/db_xref="GI:68138280"
/translation="MKKELSPHEKLLKTKQKKTKHVFIAPIFVFLMFAFMWAG
KAETPKVKTYSDDVLSAFVCGDIMGKRYVEKVTQKGADSIPOYVEPIFRASDVVAGN
FENPTVQKNTKQADKEIHLQTNKESKVLKDMNFTVLSANNHAMDYGVQGMKDTLGL
EFAKNLDIVGAGYSLSADAKKISYQKNGVGTIATLGTGDSGDKFAAKNTPGVLPA
DPEIFIPMISEAKKHADIVVQSHWGQEDNDNDROQLARMSDAGADIIVGHPH
VLEPIEVNGTVIFYSLGNFVFDQGTTRTSALVQHLKNGTGRFVTPIDTHEAT
PAVKKDSLKQKTIIRLTGKDNFAWKVEDGKLTIDHSDKLKSK"
ORIGIN
Query Match 71.1%; Score 20.2; DB 1; Length 2775;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGUATTCGACGAGCTACTCT 1810
DB 1790 CGTATTCGACGAGCTACTCT 1810
RESULT 4
AB016245 2989 bp DNA linear BCT 14-OCT-1999
DEFINITION Bacillus subtilis pgsA, pgsB and pgsC genes, complete cds.
ACCESSION AB016245
VERSION AB016245.1 GI:6045071
KEYWORDS pgsA; pgsC; pgsB.
SOURCE Bacillus subtilis
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (sites)
AUTHORS Ashiuchi,M., Soda,K. and Misono,H.
TITLE A poly-gamma-glutamate synthetic system of Bacillus subtilis IFO
3336: Gene cloning and biochemical analysis of poly-gamma-glutamate
produced by Escherichia coli clone cells
JOURNAL Biochem. Biophys. Res. Commun. 263 (1), 6-12 (1999)
PUBMED 10486244
REFERENCE 2 (bases 1 to 2989)
AUTHORS Ashiuchi,M.

```

Direct Submission
Submitted (14-JUL-1998) Makoto Ashiuchi, Kochi University, Research
Institute of Molecular Genetics; Otsu 200, Monobe, Nankoku, Kochi
783-8502, Japan (E-mail:ashiuchi@ring.kochi-u.ac.jp,
Tel:0888-64-5215, Fax:0888-64-5109)

FEATURES
source
Location/Qualifiers
1..2989
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/strain="IFO 3336"
/sub_species="Bacillus natto"
/db_xref="taxon:1423"
1..1182
/gene="pgsB"
1..1182
/gene="pgsB"
/codon_start=1
/transl_table=11
/product="pgsB"
/protein_id="BAA85263.1"
/db_xref="GI:6045072"
/translation="MWLLIACAVILVIGILEKRRHOKNIDALPVRVNINGIRKSTV
RLTFTGILIRAGYKTVGTCTDARMYWDTPPEKPIKRPQGPSIGQEKEVMEETVE
RGNAIVSECVANPDVQILIFQELLQANIGVI VNVLEDMDVNGPTLDETAFAFTAT
IPYNHLVITDSEYTEPFQKAKERNVLIADNSKITDEVLKFEYVVPDNASLAL
GVALGIDESTAFKMLNAPPDPMARILPLISFEPGFVNGFANDASSTLNINWK
RVKGIYPTDDPIIIMNCRADRVORTQOFANDVLYTBASELILIGETTPVIVKAYEE
GKIPADKLHLDLEKSTDEIMELLKKSMHNRVIYGVGNHGAAPLEIKIHEYKVKQLV
S"
1197..1646
/gene="pgsC"
1197..1646
/gene="pgsC"
/codon_start=1
/transl_table=11
/product="pgsC"
/protein_id="BAA85264.1"
/db_xref="GI:6045073"
/translation="MFGSDIYALILGLVLSLIPFAERTGIYPAGLVVPGVLGVFNQP
VFILLVLLVLLTVIVKYGLSKFMILYGRKPKAAMLTIGILVKIAPDFLYPIVPEI
APRGIGIIVPLGNTIQKGLTITFGSTLLSGATFAMFVYLI"
1665..2807
/gene="pgsA"
1665..2807
/gene="pgsA"
/codon_start=1
/transl_table=11
/product="pgsA"
/protein_id="BAA85265.1"
/db_xref="GI:6045074"
/translation="MKKELSFHEKLLKLTQKKKTKNKHVPIAIPVIFVLMFAPFMAG
KAETPKVKTSDVLSFVGDIMWGRYVEKVTQKGDSTFOYVEPIFRASDYVAGN
FENPVTQKVKQADKEIHLQTNKESVKVLKMNPTVLNANNHAMDYGVQMKDITLG
EPAKONLDIVGAGVSLDAAKKISYQKNGVTIATLGTDSVSGKPAKKNTPGLPA
DLEPIFIPMSIAKHADIIVVQSHWGQRYNDNDPDRQLARMSADAGADIIVGHHPH
VLEPIRYNGTVTFYSJGVFDQGTSTALVQYHLKKNKGTGRFVPTDIHEAT
PAPVKDSLKQKTIIRLTKDSNPAWVEDGKLTDFIDHSDKLKSK"

ORIGIN
Query Match 71.1%; Score 20.2; DB 1; Length 2989;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGUAUUCGACGACGUACUCD 21
Db 1822 CGTATTCTGACGACTACTCT 1842

RESULT 5
E50424
LOCUS E50424 3045 bp DNA linear PAT 31-JAN-2002
DEFINITION Process for producing poly-gamma-glutamic acid.
ACCESSION E50424

E50424.1 GI:18629412
JP 2001017182-A/1.
SOURCE
Bacillus subtilis
ORGANISM
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 (bases 1 to 3045)
Ashiuchi,M., Misono,H. and Soda,K.
Process for producing poly-gamma-glutamic acid
Patent: JP 2001017182-A 1 23-JAN-2001;
NAGASE & CO LTD
COMMENT
OS Bacillus subtilis (hay bacillus) IFO 3336
PN JP 2001017182-A/1
PD 23-JAN-2001
PF 09-JUL-1999 JP 1999196335
PR
PI MAKOTO ASHIUCHI,HARUO MISONO,KENJI SODA
PC C12N15/09,C08G69/36,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC
C12N9/00,C12N9/90.
PC C12P13/14,C12N15/00,C12N5/00
CC
FH Key Location/Qualifiers
FT source 1..3047
FT /organism="Bacillus subtilis"
FT /mol_type="genomic DNA"
FT /db_xref="taxon:1423"
FEATURES
source
Location/Qualifiers
1..3045
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/db_xref="taxon:1423"
ORIGIN
Query Match 71.1%; Score 20.2; DB 6; Length 3045;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGUAUUCGACGACGUACUCD 21
Db 1840 CGTATTCTGACGACTACTCT 1860

RESULT 6
AB046355
LOCUS AB046355 3738 bp DNA linear BCT 11-JAN-2002
DEFINITION Bacillus subtilis ywsc, ywtA, ywtB, ywtC genes, complete cds.
ACCESSION AB046355
VERSION AB046355.1 GI:13591556
KEYWORDS
SOURCE
Bacillus subtilis
ORGANISM
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1
Urushibata,Y., Tokuyama,S. and Tahara,Y.
Characterization of the Bacillus subtilis ywsc gene, involved in
gamma-polyglutamic acid production
J. Bacteriol. 184 (2), 337-343 (2002)
JOURNAL 11751809
PUBMED
REFERENCE 2 (bases 1 to 3738)
Tahara,Y. and Urushibata,Y.
Direct Submission
TITLE
Submitted (18-JUL-2000) Yasutaka Tahara, Shizuoka University,
Faculty of Agriculture; 836 Ohta, Shizuoka city, Shizuoka 422-8529,
Japan (E-mail:acytaha@agr.shizuoka.ac.jp,
Tel:81-54-238-4878(ex.7808), Fax:81-54-237-3028)
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
source
Location/Qualifiers
1..3738
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/db_xref="taxon:1423"
497..1678
/gene="ywsc"
497..1678
/gene="ywsc"
/codon_start=1

REFERENCE 2 (bases 1 to 4315)
AUTHORS Tran,L.P. and Itoh,Y.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2000) Yoshifumi Itoh, National Food Research
Institute, Applied Microbiology, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8642, Japan (E-mail: yosifumi@nfri.affrc.go.jp,
Tel: +81-298-38-8075, Fax: +81-298-38-7996)
FEATURES
Location/Qualifiers
1. 4315
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/strain="Asahikawa"
/db_xref="taxon:1423"
941..2122
/gene="capB"
941..2122
/gene="capB"
/codon_start=1
/transl_table=11
/product="CapB"
/protein_id="BAB13484.1"
/db_xref="GI:10119861"
/translation="MWLLIACAVILVIGLEKRRHOKNIDALPVRVNINGIRKSTV
TRLTGILIEAGYKVTGTDARMIYNDTPBEKPIKRKPGQPNIGEOKEVRETV
RGANAIVSECMANPDYQIIIFQELLOANIGVNVFEDHMDVMGPTLDEIAEAFAT
IPYNGHLVITDSYETFEFFKOKAKERTKVIADNSKITDEYLKFEVMPEDNASLAL
GVAQALGIDBEAPKGLNAPPDGAMRIILPLISPEPGHFNAGPANDASSTLIWK
RVKEIGYPTDDPIIIMNCRADVRDTCQFANDVLPYIEASELILIGETTFPIKAYEE
GKIPADKLDHLEKYSTDEIMELLKRMHNRVIYGVGNHGAEPLEIKIHEYKVKQLV
S"
2137..2586
/gene="capC"
2137..2586
/gene="capC"
/codon_start=1
/transl_table=11
/product="CapC"
/protein_id="BAB13485.1"
/db_xref="GI:10119862"
/translation="MFGSDLYIALILGVLLSLIPAEKTIIVPAGLVVPGYGLGVFNQ
VPILLVLLSVLTIVYKGLSKFMILYGRKFAAMLTIGVLKIAFDLYPIVPEFI
AEFRGIGIIVPLGIANTIQKGLTITFGSTLLSGATFAIMFVYLI"
2605..3747
/gene="capA"
2605..3747
/gene="capA"
/codon_start=1
/transl_table=11
/product="CapA"
/protein_id="BAB13486.1"
/db_xref="GI:10119863"
/translation="MKKELSFHEKLLKTKQKKTKNHVFIAPIVFVLMFPMWAG
KAPTQKTYSDVLSASFVGDIMMGYRVEKTEQKADSIQVYVEPIFRASDYVAGN
FENVTYQKYNQADKEIHLQTNKESVKVLMNFTVLSANNHAMDYGVQGMKDTLGL
EPATNLDIVGAGLSDAKKKISYQKNGVTIATLFTDVSQKGFPAKNTPGVLP
DPEIFPMISEAKHADIVVQSHGQVYNDNDNDORQLARMSDAGADIIVGHHPH
VLEPIVYNGTVIFYSIGNFVFDQGTTRDSALVQYHLKNGTGRFEVTPIDIHEAT
PAPVKDSLKQKTIIRLTQDSNFAWKVEDGKLTDFIDHSDKLKSK"
3318..3485
/gene="ywtC"
3318..3485
/gene="ywtC"
/codon_start=1
/transl_table=11
/protein_id="BAB40951.1"
/db_xref="GI:13591560"
/translation="MKFVKAIWPFVVAIVFMSAFKFNQDQTDQEKQKIDMEMNKI
QQQEPVNNK"
ORIGIN
Query Match 71.1%; Score 20.2; DB 1; Length 3738;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CGAUUCUGACGACGUACUCT 21
|||:::|||||:::|
Db 2318 CGTATCTGACGACGACTCT 2338
RESULT 7
AB039950 4315 bp DNA linear BCT 12-SEP-2000
LOCUS
DEFINITION Bacillus subtilis capB, capC, ywtC genes, complete cds.
ACCESSION AB039950
VERSION AB039950.1 GI:10119860
KEYWORDS Cap; CapC; ywtC.
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 4315)
AUTHORS Tran,L.P. and Itoh,Y.
TITLE Nucleotide sequence of the capBCA operon in Bacillus subtilis
(natto)
JOURNAL Published Only in DataBase (2000)

1 /transl_table=11
2 /protein_id="BAB40948.1"
3 /db_xref="GI:13591557"
4 /translation="MWLLIACAVILVIGLEKRRHOKNIDALPVRVNINGIRKSTV
5 TRLTGILIEAGYKVTGTDARMIYNDTPBEKPIKRKPGQPNIGEOKEVRETV
6 RGANAIVSECMANPDYQIIIFQELLOANIGVNVFEDHMDVMGPTLDEIAEAFAT
7 IPYNGHLVITDSYETFEFFKOKAKERTKVIADNSKITDEYLKFEVMPEDNASLAL
8 GVAQALGIDBEAPKGLNAPPDGAMRIILPLISPEPGHFNAGPANDASSTLIWK
9 RVKEIGYPTDDPIIIMNCRADVRDTCQFANDVLPYIEASELILIGETTFPIKAYEE
10 GKIPADKLDHLEKYSTDEIMELLKRMHNRVIYGVGNHGAEPLEIKIHEYKVKQLV
11 S"
12 1693..2142
13 /gene="ywtA"
14 /gene="ywtA"
15 1693..2142
16 /gene="ywtA"
17 /codon_start=1
18 /transl_table=11
19 /protein_id="BAB40949.1"
20 /db_xref="GI:13591558"
21 /translation="MFGSDLYIALILGVLLSLIPAEKTIIVPAGLVVPGYGLGVFNQ
22 VPILLVLLSVLTIVYKGLSKFMILYGRKFAAMLTIGVLKIAFDLYPIVPEFI
23 AEFRGIGIIVPLGIANTIQKGLTITFGSTLLSGATFAIMFVYLI"
24 2161..3303
25 /gene="ywtB"
26 /gene="ywtB"
27 2161..3303
28 /gene="ywtB"
29 /codon_start=1
30 /transl_table=11
31 /protein_id="BAB40950.1"
32 /db_xref="GI:13591559"
33 /translation="MKKELSFHEKLLKTKQKKTKNHVFIAPIVFVLMFPMWAG
34 KAPTQKTYSDVLSASFVGDIMMGYRVEKTEQKADSIQVYVEPIFRASDYVAGN
35 FENVTYQKYNQADKEIHLQTNKESVKVLMNFTVLSANNHAMDYGVQGMKDTLGL
36 EPATNLDIVGAGLSDAKKKISYQKNGVTIATLFTDVSQKGFPAKNTPGVLP
37 DPEIFPMISEAKHADIVVQSHGQVYNDNDNDORQLARMSDAGADIIVGHHPH
38 VLEPIVYNGTVIFYSIGNFVFDQGTTRDSALVQYHLKNGTGRFEVTPIDIHEAT
39 PAPVKDSLKQKTIIRLTQDSNFAWKVEDGKLTDFIDHSDKLKSK"
40 3318..3485
41 /gene="ywtC"
42 3318..3485
43 /gene="ywtC"
44 /codon_start=1
45 /transl_table=11
46 /protein_id="BAB40951.1"
47 /db_xref="GI:13591560"
48 /translation="MKFVKAIWPFVVAIVFMSAFKFNQDQTDQEKQKIDMEMNKI
49 QQQEPVNNK"
50 ORIGIN
51 Query Match 71.1%; Score 20.2; DB 1; Length 3738;
52 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
53 Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
54 Oy 1 CGAUUCUGACGACGUACUCT 21
55 |||:::|||||:::|
56 Db 2318 CGTATCTGACGACGACTCT 2338
57 RESULT 7
58 AB039950 4315 bp DNA linear BCT 12-SEP-2000
59 LOCUS
60 DEFINITION Bacillus subtilis capB, capC, ywtC genes, complete cds.
61 ACCESSION AB039950
62 VERSION AB039950.1 GI:10119860
63 KEYWORDS Cap; CapC; ywtC.
64 SOURCE Bacillus subtilis
65 ORGANISM Bacillus subtilis
66 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
67 REFERENCE 1 (bases 1 to 4315)
68 AUTHORS Tran,L.P. and Itoh,Y.
69 TITLE Nucleotide sequence of the capBCA operon in Bacillus subtilis
70 (natto)
71 JOURNAL Published Only in DataBase (2000)


```

/transl_aton="MFSGDLYALILGVLLSLIPAEKTCIVDPAGLVVPGVLGVNOL"
VFILHVLVSLTIVTVIKYGLSKFWILGRRKPAAMLITGVILVKIAPDFLYPIVPEFI
ABRFGIIGIIVPGLIANTIQKQGLITIFGSLTLLSGATAIMPFVYILI"
1538. .2694
/gene="ywtB"
/gene="ywtB"
1538. .1544
/gene="ywtB"
/gene="ywtB"
1552. .2694
/gene="ywtB"
/gene="ywtB"
/note="product highly similar to Bacillus anthracis CapA
protein"
/codon_start=1
/transl_table=11
/protein_id="CAB07469.1"
/db_xref="GI:1894767"
/db_xref="UniProt/TrEMBL:P96738"
/transl_aton="MKKELSFHEKLLKLTQKKKTNKHVFIAIPVFLVMPAFMWAG
KARTPKVKTYSQADKEIHLCTNKESVKVLKDMNFTVLNSANNHAMDVGVQMGKQTLG
EPAKONLIDVAGVYSLSDAKKISYQKGVGVTIATLGTVDVSGKGFAAKNTPGVLPAP
DPEIFIMTEAKKHHIVVQSHGQBYDNDPNRQQLARAMDAGDADIVGHHPH
VLEPIFYNGTIVPYSILGNVDFDQGTWTRDSALVOYHLKMGTRGFVTPIDIHEAT
PAPYKQSLKQKTIIRELTQDSNFAMKVEDGKLTPTDIDHSDKSLK"
2697. .2914
/gene="ywtC"
/gene="ywtC"
2697. .2702
/gene="ywtC"
2709. .2876
/gene="ywtC"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="CAB07470.1"
/db_xref="GI:1894768"
/db_xref="UniProt/Swiss-Prot:P36739"
/transl_aton="MKFVKAIWPFVAIVFMFMSAFKFNQDLTQEKQKIDWEMNKPI
QQQSEPVNANK"
2880. .2914
/gene="ywtC"
/gene="ywtC"
3000. .4258
/gene="ywtD"
/gene="ywtD"
3000. .3007
/gene="ywtD"
3017. .4258
/gene="ywtD"
/note="product similar to E. coli NLPC protein and to
Listeria species p60-related proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB07476.1"
/db_xref="GI:1894769"
/db_xref="GOA:P96740"
/db_xref="UniProt/Swiss-Prot:P96740"
/transl_aton="MNTLANWKFLNVAIICFLVPIMTKABIEADTSSELIVSEAK
NLGVQYKGGSTPEKGFDPGLIYQVPSKADILHPKSVNDQYKLGTAVKPENLKPGD
ILPEPKGSGTGVTHDHALYIGDQMVHSTQSGVLIITNYKSSYWGVTYIGARRTAA
DPATADVPVQAEAKYIGVPYVFGSGTSPSEGFDCSLGVYQVPAQALGIYLPRAEQM
KSGMTGRIRRENDLTIKPEKNTIVSEATLVYGVSPYKQGVTPETGDTAGFVQVYOKA
AGSLIPRYATYSQYNAGTKIEADLKPGDIVVFQSTSLNPSIYIGNGQVVHVTLNSGVT
ITNNWTSYWKDKVAGSIRVQ"
4259. .4293
complement(4292. .5170)
/gene="ywtE"
complement(4292. .5152)
/gene="ywtE"
/note="product similar to Bacillus subtilis YxeH and YceE
proteins and to E. coli Yida protein"
/codon_start=1
/transl_table=11
/protein_id="CAB07471.1"
/db_xref="GI:1894770"

```

```

/db_xref="GOA:P96741"
/db_xref="UniProt/TREMBL:P96741"
/translation="MKCIAIDLTGLTKNKSVISAENREAIKRAVDAGILVTICTGRA
TFDKALLDDIDIPITAAAGGTIDHTGVRLISRTLMDOEAGKAIADYLLSKNIYFEV
TDDHLSPDGEAKLHAELDLKSNAPNEQTDLLWQAGTQKQKPIPIHIESVFD
GGENYKLLCSFMDKLLQAKBELKHHKLLAQTSSGKHIIELLPASGKGATKLA
DIYGETQDIYAIGDSDNLSNFEVAGHRIANENAIDELKEKSTFVTYKSDENGVAIF
IDQLSGQYA"
complement(5163..5170)
/gene="ywtB"
5277..5305
/gene="ywtF"
5277..5305
/gene="ywtF"
5277..5305
/complement(5308..6291)
/pseudo
complement(5308..6276)
/gene="ywtF"
/notes="an in frame TGA stop codon is found at position
5599 to 5601
translation similar to Bacillus subtilis LytR protein"
/pseudo
/codon_start=1
/transl_table=11
complement(6285..6291)
/gene="ywtF"
6594..8112
/gene="ywtG"
6594..6601
/gene="ywtG"
6609..7982
/gene="ywtG"
/notes="product highly similar to metabolite transport
proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB07473.1"
/db_xref="GI:1894771"
/db_xref="GOA:P96742"
/db_xref="UniProt/TREMBL:P96742"
/translation="MKQSNILVYFPGALGGALGYDGVISGALFMKKELGLNAPT
EGLVSSLLVGAIGSGAAGKLTDRFRKKAIMAAALLFCIGGLVALAPNTGYWLVF
RIILGLAVGTSTTVPLVYSELAPKPKKGLSSNLQMLITVIGLUSYIVNTIFADAE
WRWMLGAAPVSLLLGILFMPESPRLWFTNGEESKAKILEKLGRTQIDQEIHD
KEAEKQDGGKLELDPWRPALIAGLGLAQFPQIGTNTIYYPKFTFNVFGNSA
SLTGTVIGTNNVLTVAIKTIDKIGRKLPLFGNAGWISLIVLWNLFPNNTPA
ASMTTVICLGVPIVYFVAVSGPVVWMLPELPLHVRGIGTGVSTLMHLVGLTIVSLT
YFILMEALGISVLIYIAIGIMAFLEFVRPKVTKGRSLSEIEQDLKDKNGQGAAG
KQQTVTG"
7981..8112
/gene="ywtG"
complement(8019..8189)
/gene="gerB"
complement(8019..8189)
/gene="gerB"
/codon_start=1
/transl_table=11
/product="GerB"
/protein_id="CAB07474.1"
/db_xref="GI:1894772"
/db_xref="GOA:P96743"

```

```

Query Match      71.1%; Score 20.2; DB 1; Length 8189;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGAUUUCGACGACGUACUD 21
||:::|||||:::|
Db      1709  CGTATTCTGACGACGTACTCT 1729

```

RESULT 9

```

BSUB0019/c
LOCUS
DEFINITION
  BSUB0019 200690 bp DNA linear BCT 18-APR-2005
  Bacillus subtilis complete genome (section 19 of 21) : from 3608981
  to 3809670.
ACCESSION
  Z99122 AL0009126
  Z99122.2 GI:32468830
KEYWORDS
SOURCE
  ORGANISM
    Bacillus subtilis subsp. subtilis str. 168
    Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
  1 (bases 1 to 200690)
  Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
  Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S.,
  Borries, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C.,
  Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V.,
  Carter, N.M., Choi, S.K., Codani, J.J., Conner, I.F., Cummings, N.J.,
  Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
  Emerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E.,
  Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A.,
  Galloni, N., Ghm, S.Y., Glaser, P., Goffeau, A., Gollightly, E.J.,
  Grandi, G., Guisepi, G., Guy, B.J., Haga, K., Haele, J., Harwood, C.R.,
  Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F.,
  Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,
  Klarr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P.,
  Koningsstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
  Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,
  Masuda, S., Mael, C., Medigue, C., Medina, N., Mellado, R.P.,
  Mizuno, M., Moesti, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,
  Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,
  Portetelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P.,
  Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,
  Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T.,
  Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,
  Sekowska, A., Serron, S.J., Serron, P., Shin, B.S., Soldo, B.,
  Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K.,
  Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A.,
  Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, P., Vassarotti, A.,
  Viaria, A., Wambut, R., Wedler, E., Wedler, H., Weitzenecker, T.,
  Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,
  Yoshida, K., Yoshikawa, H.F., Zumbstein, E., Yoshikawa, H. and
  Danchin, A.
  The complete genome sequence of the gram-positive bacterium
  Bacillus subtilis
  Nature 390 (6657), 249-256 (1997)
  9384377
  2 (bases 1 to 200690)
  Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
  Direct Submision
  Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,
  Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724
  Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
  adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
  68 89 48
  On Jul 7, 2003 this sequence version replaced gi:2636029.
  This entry contains data from release R16.1 of the Subtilist
  database. Further data on gene annotation and detailed information
  about changes from previous releases can be found at
  http://genolist.pasteur.fr/Subtilist/.
  Location/Qualifiers
    1..200690
    /organism="Bacillus subtilis subsp. subtilis str. 168"
    /mol_type="genomic DNA"
    /strain="168"
    /db_xref="taxon:224308"
    complement(114..2987)
    /gene="uvrA"
    /locus_tag="BSU35160"
    complement(114..2987)
    /gene="uvrA"
    /locus_tag="BSU35160"
    /function="excision of ultraviolet light-induced
    pyrimidine dimers in DNA"
    /codon_start=1

```


Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2006, 14:29:17 ; Search time 251.71 Seconds
(without alignments)
900.242 Million cell updates/sec

Title: US-10-824-829-2

Perfect score: 28.4

Sequence: 1 cguauucgacgacguacuctdcdadgdcddt 34

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

N_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002as.*
- 7: Geneseq2002bs.*
- 8: Geneseq2003as.*
- 9: Geneseq2003bs.*
- 10: Geneseq2003cs.*
- 11: Geneseq2003ds.*
- 12: Geneseq2004as.*
- 13: Geneseq2004bs.*
- 14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.2	71.1	27	13 ADT93900	Adt93900 Composite
2	20.2	71.1	1140	10 ADA44793	Ada44793 Bacillus
3	20.2	71.1	1140	12 ADO07643	Ado07643 B subtili
4	20.2	71.1	1140	13 ADV65734	Adv65734 B. subtil
5	20.2	71.1	1143	4 AAF82257	Aaf82257 Bacillus
6	20.2	71.1	3045	4 AAF82254	Aaf82254 Bacillus
7	20.2	71.1	6536	10 ADA44796	Ada44796 Microbial
8	19.6	69.0	999	4 ABL11175	Abi111175 Drosophil
9	19.6	69.0	2742	4 ABL26558	Abi26558 Drosophil
10	19.6	69.0	2999	4 ABL11174	Abi111174 Drosophil
11	18.6	65.5	42988	10 ADB74379	Adb74379 Mycobacte
12	18.6	65.5	110000	6 ABS55320_1	Continuation (2 of
13	18.6	65.5	110000	12 ADQ59398_1	Continuation (2 of
14	18.6	65.5	110000	14 ADZ13665_1	Continuation (2 of
15	18.6	65.5	341511	6 ABS55200	Abs55200 Genomic D
16	18	63.4	1419	3 ACA55827	Aac55827 S. lavend
17	18	63.4	1419	10 ADE10246	Ade10246 S. lavend
18	18	63.4	1921	4 ABL03949	Abi03949 Drosophil
19	18	63.4	4386	4 ABL03948	Abi03948 Drosophil

20	18	63.4	53500	3	AAC55842	Aac55842 Complete
21	18	63.4	53500	10	ADE10261	Ade10261 S. lavend
22	18	63.4	125910	3	AAC64370	Aac64370 Human KGN
23	17.8	62.7	750	13	ADR51234	Adr51234 Anti-biof
24	17.8	62.7	924	12	ADO80525	Ado80525 Streptomy
25	17.8	62.7	924	12	ADO80523	Ado80523 Streptomy
26	17.8	62.7	963	8	ACA39824	Aca39824 Prokaryot
27	17.8	62.7	2016	8	ADA69903	Ada69903 Rice gene
28	17.6	62.0	438	2	AAV20230	Aav20230 Probe (13
29	17.6	62.0	1038	13	ADU25587	Adu25587 cDNA enco
30	17.6	62.0	2747	4	ABL11807	Abi11807 Drosophil
31	17.6	62.0	3104	4	ABL25812	Abi25812 Drosophil
32	17.6	62.0	3487	14	ADW44491	Adw44491 Zebrafish
33	17.6	62.0	5337	4	ABL11806	Abi11806 Drosophil
34	17.4	61.3	128	8	ACA13632	Aca13632 Prokaryot
35	17.4	61.3	388	8	ABX42554	Abx42554 Bovine ES
36	17.4	61.3	440	3	AAF08228	Aaf08228 Fusarium
37	17.4	61.3	440	13	ADU52269	Adu52269 Fusarium
38	17.4	61.3	440	14	ADZ90272	Adz90272 Fusarium
39	17.4	61.3	1154	2	AAZ20170	Aaz20170 Bovine pr
40	17.4	61.3	1154	10	ADF11576	Adf11576 Bovine pr
41	17.4	61.3	1239	8	ACA26709	Aca26709 Prokaryot
42	17.4	61.3	1242	8	ACA24087	Aca24087 Prokaryot
43	17.4	61.3	83120	9	RAAL57571	Ral57571 Human CGI
44	17.4	61.3	296405	11	ACN44538	Acn44538 Human Gen
45	17.2	60.6	433	5	AAS70723	Aas70723 DNA encod

ALIGNMENTS

RESULT 1

ID ADT93900 standard; DNA; 27 BP.

XX AC ADT93900;

XX DT 13-JAN-2005 (first entry)

XX Composite primer oligonucleotide #2 for global amplification method.

XX ss; primer; amplification; DNA-dependent DNA polymerase;

KW RNA-dependent DNA polymerase; sequence analysis; polymorphism detection;

KW oligonucleotide ligation-based assay; comparative genome hybridization;

KW expression profiling; multiple genotype determination;

KW subtractive hybridization probe; array based assay;

KW global amplification reaction.

XX OS Synthetic.

XX Key Location/Qualifiers

FT misc_RNA 1..20

FT /*tag= a

XX WO2004092418-A2.

XX PD 28-OCT-2004.

XX PF 14-APR-2004; 2004WO-US012779.

XX PR 14-APR-2003; 2003US-0462962P.

XX PR 14-APR-2003; 2003US-0462965P.

XX PA (NUGE-) NUGEN TECHNOLOGIES INC.

XX Kurn N, Wang S;

XX WPI; 2004-784618/77.

XX Amplifying a template polynucleotide useful in expression profiling, by incubating a template polynucleotide, a composite primer comprising RNA and 3' DNA portion, a DNA-dependent DNA polymerase, and RNA-dependent DNA polymerase.

XX PS Claim 41; SEQ ID NO 2; 95pp; English.

XX CC The invention relates to a method of amplifying a template polynucleotide

XX CC by incubating a reaction mixture comprising a template polynucleotide, a

XX CC composite primer that is hybridizable to template polynucleotide sites

XX CC and comprises an RNA portion and a 3' DNA portion, a DNA-dependent DNA

XX CC polymerase, and an RNA-dependent DNA polymerase, for first primer random

XX CC hybridization to the template polynucleotide, and primer extension, where

XX CC RNA/DNA heteroduplex is generated. The method is useful for amplifying a

XX CC template polynucleotide which is useful for further analysis such as

XX CC sequence analysis, polymorphism detection using e.g. oligonucleotide

XX CC ligation-based assays, analysis using invader, cleavage or limited primer

XX CC extension. The method or the kit is useful for pre-amplifying a pool of

XX CC polynucleotides to generate sufficient input material for subsequent

XX CC analysis. The method is useful for generating polynucleotide products

XX CC used in comparative genome hybridization, expression profiling, multiple

XX CC genotype determinations, generation of cDNA libraries, generation of

XX CC subtractive hybridization probes, and array based assays. This sequence

XX CC represents a composite primer used to perform global amplification

XX CC reactions on human genomic DNA as a template.

XX SQ Sequence 27 BP; 5 A; 9 C; 5 G; 2 T; 6 U; 0 Other;

Query Match 71.1%; Score 20.2; DB 13; Length 27;

Best Local Similarity 95.2%; Pred. No. 7.5;

Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGUAUUCUGACGACGUACUCD 21

Db 1 CGUAUUCUGACGACGUACUC 21

RESULT 2

ADA44793

ID ADA44793 standard; DNA; 1140 BP.

XX AC ADA44793;

XX DT 20-NOV-2003 (first entry)

XX DE Bacillus subtilis poly-gamma-glutamate synthetase complex pgSA gene.

XX KW Cell surface; expression vector; microbial;

XX KW poly-gamma-glutamate synthetase; pgBCA complex; surface expression;

XX KW Gram-positive bacterium; Gram-negative bacterium; enzyme; antigen;

XX KW antibody; attachment protein; adsorption protein; vaccine; pgSA; gene;

XX KW ds.

XX OS Bacillus subtilis; variety chungkookjang.

XX PN WO2003014360-A1.

XX PD 20-FEB-2003.

XX PF 09-AUG-2002; 2002WO-KR001522.

XX PR 10-AUG-2001; 2001KR-00048373.

XX PA (BIOL-) BIOLEADERS CORP.

XX PA (MDMD-) MD LAB CO LTD.

XX PI Sung M, Hong S, Lee J, Jung C, Kim C, Soda K, Ashiuchi M;

XX DR WPI; 2003-256589/25.

XX PT New expression vector containing gene(s) that encode a poly-gamma-

XX PT glutamate synthetase complex, useful for producing proteins (e.g.

XX PT vaccines or enzymes) on the microbial surface of Gram-positive and/or

XX PT Gram-negative bacteria.

XX PS Claim 3; Page 109-110; 122pp; English.

CC The invention relates to a vector for expression of a target protein on a

CC microbial cell surface. The vector of the invention comprises either one

CC or more than two genes of the Bacillus subtilis poly-gamma-glutamate

CC synthetase complex (pgBCA) to facilitate microbial surface expression of

CC the target protein. The pgBCA gene complex comprises the pgSB, pgSB and

CC pgSA genes and is normally expressed in the outer membrane of Bacillus

CC subtilis. The vector can be transformed into either Gram-positive or Gram

CC -negative bacteria (e.g., Escherichia coli), and can be used for the

CC surface expression of various proteins of interest such as enzymes,

CC antigens, antibodies, attachment proteins or adsorption proteins.

CC Proteins recombinantly produced using the vector of the invention can be

CC used as, for example, vaccines or enzymes. The present sequence

CC represents the Bacillus subtilis pgBCA complex gene pgSA, which is

CC specifically claimed for use in the vector of the invention.

XX SQ Sequence 1140 BP; 381 A; 235 C; 265 G; 259 T; 0 U; 0 Other;

Query Match 71.1%; Score 20.2; DB 10; Length 1140;

Best Local Similarity 66.7%; Pred. No. 16;

Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGUAUUCUGACGACGUACUCD 21

Db 158 CGTATTCTGACGACGTACTCT 178

RESULT 3

ADO07643

ID ADO07643 standard; DNA; 1140 BP.

XX AC ADO07643;

XX DT 15-JUL-2004 (first entry)

XX DE B subtilis poly-X-glutamate synthetase complex coding sequence pgSA.

XX KW ds; gene; enzyme; vaccine; cytosstatic; pgSA; poly-X-glutamate synthetase;

XX KW human papilloma virus.

XX OS Bacillus subtilis.

XX PN WO2004035795-A1.

XX PD 29-APR-2004.

XX PF 17-OCT-2003; 2003WO-KR002163.

XX PR 17-OCT-2002; 2002KR-00063378.

XX PA (BIOL-) BIOLEADERS CORP.

XX PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.

XX PI Sung M, Poo H, Lee J, Jung C, Hong S, Kim C, Park S, Pyo H;

XX DR WPI; 2004-348463/32.

XX PT New vector containing pgs A-C genes encoding poly-gamma-glutamate

XX PT synthetase complex and an antigen protein gene of human papilloma virus,

XX PT useful in preparing vaccine for treating or preventing mucosal tumor,

XX PT e.g. cervical cancer.

XX PS Disclosure; Page 62-63; 69pp; English.

XX CC The present invention relates to a vector for preparing a vaccine which

XX CC contains one or more than two genes, i.e. pgs A-C encoding poly-X-

XX CC glutamate synthetase complex and an antigen protein gene of human

XX CC papilloma virus. The vector and microbes transformed with it are useful

XX CC in preparing vaccines for treating or preventing mucosal tumour, e.g.

XX CC cervical cancer. The present sequence is a Bacillus subtilis poly-X-

XX CC glutamate synthetase complex coding sequence.

XX SQ Sequence 1140 BP; 381 A; 235 C; 265 G; 259 T; 0 U; 0 Other;

Query Match 71.1%; Score 20.2; DB 12; Length 1140;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGUAAUCUGACGACGUACUCD 21
 Db 158 CGTATTCGACGACGTACTCT 178

RESULT 4

ADV65734
 ID ADV65734 standard; DNA; 1140 BP.

XX ADV65734;

XX 10-FEB-2005 (first entry)

DE B. subtilis poly-gamma-glutamic acid synthase A gene SEQ ID NO:3.

XX ds; poly-gamma-glutamic acid synthase A; pgsA; gene expression;
 KW antibiotic; antimicrobial; fungicide; cytostatic.

XX Bacillus subtilis.

XX KR2004034780-A.

XX 29-APR-2004.

XX 17-OCT-2002; 2002KR-00063379.

XX 17-OCT-2002; 2002KR-00063379.

XX (BIOL-) BIOLEADERS CORP.

PA (UYCH-) UNIV CHUNSON CO LTD.

PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.

XX Boo HR, Ham GS, Hong SP, Jung CM, Kim CU, Lee DG, Lee JS;

PI Park YG, Sung MH;

XX WPI; 2004-577380/56.

XX Method for surface expression of peptides p5 and anal3 using pgs bca
 gene.

XX Disclosure; SEQ ID NO 3; 25pp; Korean.

CC The invention relates to a novel method for surface expression of
 CC peptides P5 and Anal3 using a poly-gamma-glutamic acid synthase (pgs) bca
 CC gene, thereby removing a purification process of peptides P5 and Anal3,
 CC and using lactic acid bacteria for the surface expression, so that
 CC peptide antibiotics can be cheaply and stably mass-produced. An
 CC expression vector pHEILB:pgsA-P5 comprises one or more genes encoding
 CC poly-gamma-glutamic acid synthase selected from pgsB, pgsC and pgsA, and
 CC a gene for dipolar peptide antibiotics having antimicrobial, antifungal
 CC and anticancer activities, wherein the dipolar peptide antibiotic has
 CC homology to the peptide P5 encoded by the nucleotide sequence set forth
 CC in ADV65735, or to the peptide Anal3 encoded by the nucleotide sequence
 CC set forth in ADV65737. The present sequence represents the B. subtilis
 CC poly-gamma-glutamic acid synthase A gene used in the invention.

XX Sequence 1140 BP; 381 A; 235 C; 265 G; 259 T; 0 U; 0 Other;

Query Match 71.1%; Score 20.2; DB 13; Length 1140;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGUAAUCUGACGACGUACUCD 21
 Db 158 CGTATTCGACGACGTACTCT 178

RESULT 5

AAF82257

ID AAF82257 standard; DNA; 1143 BP.

XX AAF82257;

XX 21-JUN-2001 (first entry)

DE Bacillus subtilis IFO 3336 DNA encoding a pga synthesising enzyme.

XX Bacterial; poly-gamma-L-glutamic acid; L-PGA;

KW poly-gamma-D,L-glutamic acid; PGA; glutamate racemase; ds.

XX Bacillus subtilis.

XX Key Location/Qualifiers

FT 1. .1143

FT /*tag= a

XX JP2001017182-A.

XX 23-JAN-2001.

XX 09-JUL-1999; 99JP-00196335.

XX 09-JUL-1999; 99JP-00196335.

XX (NAGS) NAGASE SANGYO KK.

XX WPI; 2001-285408/30.

DR P-PSDB; AAB74026.

XX New nucleic acid encoding a glutamate racemase enzyme useful for the
 PT preparation of poly-gamma-glutamic acid.

XX Claim 4; Page 13-15; 17pp; Japanese.

CC The present sequence encodes an enzyme which is useful in the production
 CC of poly-gamma-L-glutamic acid (L-PGA) or poly-gamma-D,L-glutamic acid
 CC (PGA). A plasmid comprising the present sequence may be used to transform
 CC Escherichia coli. The transformants express the enzyme and PGA is
 CC produced in the culture

XX Sequence 1143 BP; 385 A; 233 C; 264 G; 261 T; 0 U; 0 Other;

Query Match 71.1%; Score 20.2; DB 4; Length 1143;
 Best Local Similarity 66.7%; Pred. No. 16;

Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGUAAUCUGACGACGUACUCD 21

Db 158 CGTATTCGACGACGTACTCT 178

RESULT 6

AAF82254

ID AAF82254 standard; DNA; 3045 BP.

XX AAF82254;

XX 21-JUN-2001 (first entry)

DE Bacillus subtilis IFO 3336 DNA encoding glutamate racemase enzyme.

XX Bacterial; poly-gamma-L-glutamic acid; L-PGA;

KW poly-gamma-D,L-glutamic acid; PGA; glutamate racemase; ds.

XX Bacillus subtilis.

XX Key Location/Qualifiers

FT 19. .1200

FT /*tag= a

FT 1215. .1264

FT /*tag= b

FT 1683. .2825

```

PT      . . . . . /*tag= c
XX      JP2001017182-A.
XX      23-JAN-2001.
XX      09-JUL-1999; 99JP-00196335.
XX      09-JUL-1999; 99JP-00196335.
XX      (NAGS ) NAGASE SANGYO KK.
XX      WPI; 2001-285408/30.
XX      P-PSDB; AAB74024, AAB74025, AAB74026.
XX      New nucleic acid encoding a glutamate racemase enzyme useful for the
XX      preparation of poly-gamma-glutamic acid.
XX      Claim 1; Page 7-11; 17pp; Japanese.
XX      The present sequence encodes a glutamate racemase enzyme which is useful
XX      in the production of poly-gamma-L-glutamic acid (L-PGA) or poly-gamma-D,L
XX      glutamic acid (PGA). A plasmid comprising the present sequence may be
XX      used to transform Escherichia coli. The transformants express the enzyme
XX      and PGA is produced in the culture
XX      Sequence 3045 BP; 1000 A; 593 C; 696 G; 756 T; 0 U; 0 Other;
XX      Query Match 71.1%; Score 20.2; DB 4; Length 3045;
XX      Best Local Similarity 66.7%; Pred. No. 19;
XX      Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
XX
QY      1 CGUAUUCUGACGACGUACUCD 21
DB      1840 CGTATTCTGACGACGTACTCT 1860
XX
RESULT 7
ADA44796
ID      ADA44796 standard; DNA; 6536 BP.
XX
AC      ADA44796;
XX
DT      20-NOV-2003 (first entry)
XX
DE      Microbial cell surface expression vector pGNEBCA, SEQ ID NO:6.
XX
KW      Cell surface; expression vector; microbial;
XX      poly-gamma-glutamate synthetase; Bacillus subtilis pgsBCA complex;
XX      surface expression; Gram-positive bacterium; Gram-negative bacterium;
XX      enzyme; antigen; antibody; attachment protein; adsorption protein;
XX      vaccine; pGNEBCA; cyclic; circular; ds.
XX
OS      Synthetic.
XX      Bacillus subtilis; variety chungkookjang.
XX
PN      WO2003014360-A1.
XX
PD      20-FEB-2003.
XX
PF      09-AUG-2002; 2002WO-KR001522.
XX
PR      10-AUG-2001; 2001KR-00048373.
XX
PA      (BIOL-) BIOLEADERS CORP.
XX      (MDMD-) MD LAB CO LTD.
XX
PI      Sung M, Hong S, Lee J, Jung C, Kim C, Soda K, Ashiuchi M;
XX      WPI; 2003-256589/25.
XX      New expression vector containing gene(s) that encode a poly-gamma-
XX      glutamate synthetase complex, useful for producing proteins (e.g.

```

```

PT      vaccines or enzymes) on the microbial surface of Gram-positive and/or
XX      Gram-negative bacteria.
XX      Example 1; Page 113-120; 122pp; English.
XX      The invention relates to a vector for expression of a target protein on a
XX      microbial cell surface. The vector of the invention comprises either one
XX      or more than two genes of the Bacillus subtilis poly-gamma-glutamate
XX      synthetase complex (pgsBCA) to facilitate microbial surface expression of
XX      the target protein. The pgsBCA gene complex comprises the pgsB, pgsS and
XX      pgsA genes and is normally expressed in the outer membrane of Bacillus
XX      subtilis. The vector can be transformed into either Gram-positive or Gram
XX      negative bacteria (e.g., Escherichia coli), and can be used for the
XX      surface expression of various proteins of interest such as enzymes,
XX      antigens, antibodies, attachment proteins or adsorption proteins.
XX      Proteins recombinantly produced using the vector of the invention can be
XX      used as, for example, vaccines or enzymes. The present sequence
XX      represents a microbial cell surface expression vector of the invention,
XX      pGNEBCA.
XX      Sequence 6536 BP; 1855 A; 1464 C; 1567 G; 1649 T; 0 U; 1 Other;
XX      Query Match 71.1%; Score 20.2; DB 10; Length 6536;
XX      Best Local Similarity 66.7%; Pred. No. 22;
XX      Matches 14; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
XX
QY      1 CGUAUUCUGACGACGUACUCD 21
DB      2048 CGTATTCTGACGACGTACTCT 2068
XX
RESULT 8
ABL11175/c
ID      ABL11175 standard; cDNA; 999 BP.
XX
AC      ABL11175;
XX
DT      26-MAR-2002 (first entry)
XX
DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 28007.
XX
KW      Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical; gene; ss.
XX
OS      Drosophila melanogaster.
XX
PN      WO200171042-A2.
XX
PD      27-SEP-2001.
XX
PF      23-MAR-2001; 2001WO-US009231.
XX
PR      23-MAR-2000; 2000US-0191637P.
XX      11-JUL-2000; 2000US-00614150.
XX
PA      (PEKE ) PE CORP NY.
XX
PI      Venter JC, Adams M, Li PWD, Myers EW;
XX      WPI; 2001-656860/75.
XX      P-PSDB; ABB67072.
XX
PT      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions.
XX      Claim 1; SEQ ID NO 28007; 21pp + Sequence Listing; English.
XX      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention

```

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 999 BP; 273 A; 232 C; 259 G; 235 T; 0 U; 0 Other;

Query Match 69.0%; Score 19.6; DB 4; Length 999;

Best Local Similarity 52.9%; Pred. No. 30;

Matches 18; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGUAUUCGACGACGUACUCUPTDCDADGDCDCDT 34

DB 218 CGGATTCTCAGCAGCATTCATCGCGGCGTGCAT 185

RESULT 9

ABL26558/c

ID ABL26558 standard; DNA; 2742 BP;

XX ABL26558;

AC ABL26558;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 31147.

DB Drosophila melanogaster genomic polynucleotide SEQ ID NO 31147.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

OS Drosophila melanogaster.

XX WO200171042-A2.

FN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

XX Claim 1; SEQ ID NO 31147; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2742 BP; 688 A; 704 C; 695 G; 655 T; 0 U; 0 Other;

Query Match 69.0%; Score 19.6; DB 4; Length 2742;

Best Local Similarity 52.9%; Pred. No. 36;

Matches 18; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGUAUUCGACGACGUACUCUPTDCDADGDCDCDT 34

DB 117 CGGATTCTCAGCAGCATTCATCGCGGCGTGCAT 84

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 999 BP; 273 A; 232 C; 259 G; 235 T; 0 U; 0 Other;

Query Match 69.0%; Score 19.6; DB 4; Length 999;

Best Local Similarity 52.9%; Pred. No. 30;

Matches 18; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGUAUUCGACGACGUACUCUPTDCDADGDCDCDT 34

DB 218 CGGATTCTCAGCAGCATTCATCGCGGCGTGCAT 185

RESULT 9

ABL26558/c

ID ABL26558 standard; DNA; 2742 BP;

XX ABL26558;

AC ABL26558;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 31147.

DB Drosophila melanogaster genomic polynucleotide SEQ ID NO 31147.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

OS Drosophila melanogaster.

XX WO200171042-A2.

FN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

XX Claim 1; SEQ ID NO 31147; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2742 BP; 688 A; 704 C; 695 G; 655 T; 0 U; 0 Other;

Query Match 69.0%; Score 19.6; DB 4; Length 2742;

Best Local Similarity 52.9%; Pred. No. 36;

Matches 18; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGUAUUCGACGACGUACUCUPTDCDADGDCDCDT 34

DB 117 CGGATTCTCAGCAGCATTCATCGCGGCGTGCAT 84

RESULT 10

ABL11174/c

ID ABL11174 standard; cDNA; 2999 BP.

XX ABL11174;

AC ABL11174;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 28004.

DE Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

XX Claim 1; SEQ ID NO 28004; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2999 BP; 835 A; 688 C; 705 G; 771 T; 0 U; 0 Other;

Query Match 69.0%; Score 19.6; DB 4; Length 2999;

Best Local Similarity 52.9%; Pred. No. 37;

Matches 18; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGUAUUCGACGACGUACUCUPTDCDADGDCDCDT 34

DB 1218 CGGATTCTCAGCAGCATTCATCGCGGCGTGCAT 1185

RESULT 11

ADB74379/c

ID ADB74379 standard; DNA; 42988 BP.

XX ADB74379;

XX 04-DEC-2003 (first entry)

XX Mycobacterium leprae DNA #13.

XX Non-naturally occurring peptide; anion pump protein; tuberculosis;

KW hypersensitivity reaction; tuberculostatic; gene; ds.

XX OS Mycobacterium leprae.
XX XX
PN US6583266-B1.
XX XX
PD 24-JUN-2003.
XX XX
XX 16-SEP-1994; 94US-003111731.
XX XX
XX 19-AUG-1993; 93US-00109181.
PR 22-OCT-1993; 93US-00142558.
XX XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX PA
XX PI Smith DR, Mao J;
XX XX
XX WPI; 2003-656441/62.
XX XX
XX New Mycobacterium tuberculosis anion pump peptide useful for as
PT tuberculosis vaccine and diagnosis of tuberculosis infection.
XX XX
XX Disclosure; SEQ ID NO 128; 26pp; English.
XX XX
XX The invention relates to a non-naturally occurring peptide of
CC Mycobacterium tuberculosis comprising an amino acid sequence
CC corresponding to an anion pump protein. The invention also relates to a
CC non-naturally occurring nucleic acid corresponding to a DNA sequence of
CC Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is
CC useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium
CC leprae or for screening for new tuberculosis drugs. Purified proteins
CC derived from the sequences of the invention may elicit a specific immune
CC response. The peptide may also be used to detect hypersensitivity or
CC reactions of individuals exposed to Mycobacterium tuberculosis or
CC Mycobacterium leprae. The proteins and peptides may be affixed to solid
CC supports to detect antibodies typical of hypersensitivity reactions, from
CC a patient's sera. This sequence represents Mycobacterium leprae DNA of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html.
XX XX
XX Sequence 42988 BP; 9586 A; 13687 C; 11946 G; 7768 T; 0 U; 1 Other;
SQ
Query Match 65.5%; Score 18.6; DB 10; Length 42988;
Best Local Similarity 48.5%; Pred. No. 1.9e+02;
Matches 16; Conservative 11; Mismatches 6; Indels 0; Gaps 0;
QY 2 GUUUUCUGACGACGUACUCDTDCDADGDCDCDT 34
Db 8776 GTGTTCTGTGAAGCACTCTTTGGAAGGCGCCT 8744
RESULT 12
ABS55320_1/c
Continuation (2 of 5) of ABS55320 from base 100001 (Human transporter protein genomic
WP Sequence split into 5 fragments LOCUS ABS55320 Accession ABS55320
WP Fragment Name Begin End
WP ABS55320_0 1 110000
WP ABS55320_1 100001 210000
WP ABS55320_2 200001 310000
WP ABS55320_3 300001 410000
WP ABS55320_4 400001 506286
Query Match 65.5%; Score 18.6; DB 6; Length 110000;
Best Local Similarity 51.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 10; Mismatches 6; Indels 0; Gaps 0;
QY 2 GUUUUCUGACGACGUACUCDTDCDADGDCDCDT 34
Db 29343 GCACTCTGACGTGGTACTCTTTCACTGACCCAT 29311
RESULT 13
ADQ59398_1/c

Continuation (2 of 4) of ADQ59398 from base 100001 (Human cancer-associated (CA) gene seq
WP Sequence split into 4 fragments LOCUS ADQ59398 Accession ADQ59398
WP Fragment Name Begin End
WP ADQ59398_0 1 110000
WP ADQ59398_1 100001 210000
WP ADQ59398_2 200001 310000
WP ADQ59398_3 300001 357652
Query Match 65.5%; Score 18.6; DB 12; Length 110000;
Best Local Similarity 51.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 10; Mismatches 6; Indels 0; Gaps 0;
QY 2 GUUUUCUGACGACGUACUCDTDCDADGDCDCDT 34
Db 37505 GCACTCTGACGTGGTACTCTTTCACTGACCCAT 37473
RESULT 14
ADZ13665_1/c
Continuation (2 of 5) of ADZ13665 from base 100001 (Human cancer-associated genomic DNA #
WP Sequence split into 5 fragments LOCUS ADZ13665 Accession ADZ13665
WP Fragment Name Begin End
WP ADZ13665_0 1 110000
WP ADZ13665_1 100001 210000
WP ADZ13665_2 200001 310000
WP ADZ13665_3 300001 410000
WP ADZ13665_4 400001 443039
Query Match 65.5%; Score 18.6; DB 14; Length 110000;
Best Local Similarity 51.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 10; Mismatches 6; Indels 0; Gaps 0;
QY 2 GUUUUCUGACGACGUACUCDTDCDADGDCDCDT 34
Db 37502 GCACTCTGACGTGGTACTCTTTCACTGACCCAT 37470
RESULT 15
ABS55200/c
ID ABS55200 standard; DNA; 341511 BP.
XX AC ABS55200;
XX XX
DT 06-JAN-2003 (first entry)
XX XX
DE Genomic DNA encoding human transporter protein.
XX XX
KW Human; transporter protein; hypoglycaemia; antidiabetic; gene therapy;
KW gene; ds; single nucleotide polymorphism; SNP.
XX OS Homo sapiens.
XX XX
FH Key Location/Qualifiers
FT CDS 2001..339512
FT /tag= a
FT /product= "Human transporter protein"
FT /number= 1
FT /tag= b
FT /number= 1
FT /tag= c
FT /number= 1
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /replace(5492,C)
FT /tag= x
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /replace(5930,A)
FT /tag= y
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /replace(7269,C)
FT /tag= z

FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	variation	replace(29244,C)	FT	variation	replace(29244,C)
FT	variation	/tag= aa	FT	variation	/tag= ay	FT	variation	/tag= ay
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	variation	replace(29651,T)	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(10664,T)	FT	variation	/tag= az	FT	variation	/tag= az
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	exon	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	exon	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(11043,A/G/C)	FT	exon	30238. .30361	FT	exon	30238. .30361
FT	variation	/tag= ac	FT	exon	/tag= d	FT	exon	/tag= d
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/number= 2	FT	intron	/number= 2
FT	variation	replace(13109,T)	FT	intron	30362. .40719	FT	intron	30362. .40719
FT	variation	/tag= ad	FT	intron	/tag= e	FT	intron	/tag= e
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/number= 2	FT	intron	/number= 2
FT	variation	replace(13392,A)	FT	intron	replace(34034,T)	FT	intron	replace(34034,T)
FT	variation	/tag= ae	FT	intron	/tag= ba	FT	intron	/tag= ba
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(13395. .13397,GA)	FT	intron	replace(36671,G)	FT	intron	replace(36671,G)
FT	variation	/tag= af	FT	intron	/tag= bb	FT	intron	/tag= bb
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(13563,C)	FT	intron	replace(36836,T)	FT	intron	replace(36836,T)
FT	variation	/tag= ag	FT	intron	/tag= bc	FT	intron	/tag= bc
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(13694,A)	FT	intron	replace(37010,A)	FT	intron	replace(37010,A)
FT	variation	/tag= ah	FT	intron	/tag= bd	FT	intron	/tag= bd
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(14075,A)	FT	intron	replace(38461,A/C)	FT	intron	replace(38461,A/C)
FT	variation	/tag= ai	FT	intron	/tag= be	FT	intron	/tag= be
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(14573,A)	FT	intron	replace(38588,A/C)	FT	intron	replace(38588,A/C)
FT	variation	/tag= aj	FT	intron	/tag= bf	FT	intron	/tag= bf
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(15013,T)	FT	intron	replace(39673,A)	FT	intron	replace(39673,A)
FT	variation	/tag= ak	FT	intron	/tag= bg	FT	intron	/tag= bg
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(15073. .15075,GT)	FT	intron	replace(39854,C)	FT	intron	replace(39854,C)
FT	variation	/tag= al	FT	intron	/tag= bh	FT	intron	/tag= bh
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(15848,T)	FT	intron	40720. .40826	FT	intron	40720. .40826
FT	variation	/tag= am	FT	intron	/tag= f	FT	intron	/tag= f
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/number= 3	FT	intron	/number= 3
FT	variation	replace(16346,T)	FT	intron	40827. .45732	FT	intron	40827. .45732
FT	variation	/tag= an	FT	intron	/tag= g	FT	intron	/tag= g
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/number= 3	FT	intron	/number= 3
FT	variation	replace(16367,A)	FT	intron	replace(41325,A)	FT	intron	replace(41325,A)
FT	variation	/tag= ao	FT	intron	/tag= bi	FT	intron	/tag= bi
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(16639,G)	FT	intron	replace(41412,A)	FT	intron	replace(41412,A)
FT	variation	/tag= ap	FT	intron	/tag= bj	FT	intron	/tag= bj
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(18149,A/T)	FT	intron	replace(41475,G)	FT	intron	replace(41475,G)
FT	variation	/tag= aq	FT	intron	/tag= bk	FT	intron	/tag= bk
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(18243,C/A)	FT	intron	replace(41541,G)	FT	intron	replace(41541,G)
FT	variation	/tag= ar	FT	intron	/tag= bl	FT	intron	/tag= bl
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(18293,G/A)	FT	intron	replace(41649,G)	FT	intron	replace(41649,G)
FT	variation	/tag= as	FT	intron	/tag= bm	FT	intron	/tag= bm
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(19185,C)	FT	intron	replace(41887,T)	FT	intron	replace(41887,T)
FT	variation	/tag= at	FT	intron	/tag= bn	FT	intron	/tag= bn
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(24728. .24730,AA)	FT	intron	replace(42790,A)	FT	intron	replace(42790,A)
FT	variation	/tag= au	FT	intron	/tag= bo	FT	intron	/tag= bo
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	replace(24790,T)	FT	intron	45733. .45861	FT	intron	45733. .45861
FT	variation	/tag= av	FT	intron	/tag= h	FT	intron	/tag= h
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/number= 4	FT	intron	/number= 4
FT	variation	replace(24950,T)	FT	intron	45862. .59818	FT	intron	45862. .59818
FT	variation	/tag= aw	FT	intron	/tag= i	FT	intron	/tag= i
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/number= 4	FT	intron	/number= 4
FT	variation	replace(28608,G)	FT	intron	replace(46178,T)	FT	intron	replace(46178,T)
FT	variation	/tag= ax	FT	intron	/tag= bp	FT	intron	/tag= bp
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "	FT	intron	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation		FT	intron	replace(46282,C)	FT	intron	replace(46282,C)

```

FT      /*tag= bq
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(46382,C)
FT      /*tag= br
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(46550,A)
FT      /*tag= bs
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(49271,.49273,AT)
FT      /*tag= bt
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(49586,A)
FT      /*tag= bu
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(49835,G)
FT      /*tag= bv
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(52080,G)
FT      /*tag= bw
FT      /standard_name= "Single nucleotide polymorphism (SNP) "

```

```

Query Match      65.5%; Score 18.6; DB 6; Length 341511;
Best Local Similarity 51.5%; Pred. No. 2.9e+02;
Matches 17; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      2  GUUUCUGACGACGACUCUCDCDADGDCDCDT 34
Db      129343  GCACCTCGACGTGGTACTTTTCACTGACCCAT 129311

```

```

Search completed: February 18, 2006, 14:45:35
Job time : 254.71 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 18, 2006, 14:45:52 ; Search time 2102.52 Seconds
(without alignments)
756.599 Million cell updates/sec

Title: US-10-824-829-2
Perfect score: 28.4
Sequence: 1 gbaucugacgacguacucdcdadgdcdcdt 34

Scoring table: IDENTITY.NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

- EST:
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.4	71.8	1087	9	CC277488
2	20.2	71.1	385	3	BP096293
3	20	70.4	401	9	AQ909011
4	20	70.4	411	9	AQ444240
5	19.6	69.0	383	2	BF488748
6	19.6	69.0	531	2	BF487051
7	19.6	69.0	560	2	BF490222
8	19.6	69.0	648	2	BF503585
9	19.6	69.0	653	2	BF499434
10	19.6	69.0	703	2	BF486138
11	19.4	68.3	168	10	EX653933
12	19.4	68.3	370	6	CF280695
13	19.4	68.3	428	6	CF276319
14	19.4	68.3	444	6	CF311457
15	19.4	68.3	477	7	CO335177
16	19.4	68.3	579	7	CK082523
17	19.4	68.3	659	6	CA486252
18	19.4	68.3	665	7	CF958197
19	19.4	68.3	675	1	AU088604
20	19.4	68.3	682	1	AU093418
21	19.4	68.3	726	5	C97999
22	19.4	68.3	776	6	CB684460

23	19.4	68.3	915	10	CL481572	
24	19.4	68.3	1035	2	BG706747	
25	19.2	67.6	450	9	B04878	
26	19	66.9	477	10	CM033701	
c	27	19	562	7	CM464486	
28	19	66.9	582	6	CD180886	
c	29	19	66.9	650	3	BI935096
c	30	19	66.9	706	3	BI924562
c	31	19	66.9	727	3	BI934841
c	32	19	66.9	736	3	BI766977
c	33	19	66.9	829	8	DN940110
c	34	19	66.9	831	8	DR912306
c	35	19	66.9	869	8	DR929463
c	36	19	66.9	935	7	CK275265
c	37	18.8	66.2	337	9	BH003882
c	38	18.8	66.2	727	1	AJ934074
c	39	18.8	66.2	783	1	AJ921675
c	40	18.8	66.2	804	1	AJ924119
c	41	18.8	66.2	805	1	AJ922655
c	42	18.8	66.2	810	1	AJ921152
c	43	18.8	66.2	818	1	AJ923047
c	44	18.8	66.2	845	9	AQ891901
c	45	18.8	66.2	874	10	CG339565

ALIGNMENTS

RESULT 1
CC277488 1087 bp DNA linear GSS 13-MAY-2003
LOCUS CH261-22M6_Sp6.1 CH261 Gallus gallus genomic clone CH261-22M6,
DEFINITION genomic survey sequence.
ACCESSION CC277488
VERSION CC277488.1 GI:30637722
KEYWORDS GSS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Sukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1087)
Kremizki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
CONTACT: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 20
High quality sequence stop: 556.
Location/Qualifiers
1. 1087
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-22M6"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"

FEATURES

source
1. 1087
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-22M6"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 71.8%; Score 20.4; DB 9; Length 1087;
Best Local Similarity 53.3%; Pred. No. 1.1e+02;
Matches 16; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

COMMENT

Contact: Sanchez D.O.
Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral San Martin)
Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos Aires, Argentina
Tel: (54-11) 4580/7255/7
Fax: (54-11) 4752-9639
Email: dsanchez@ib.unsam.edu.ar
Seq primer: T7
Class: shotgun.

FEATURES

source	Location/Qualifiers
	1..401
	/organism="Trypanosoma cruzi"
	/mol_type="genomic DNA"
	/strain="CL-Brener"
	/db_xref="taxon:5693"
	/clone="G4B6"
	/cell_type="epimastigote"
	/clone_lib="Trypanosoma cruzi random genomic library"
	/note="Vector: pBS(-) (Stratagene); T. cruzi DNA was randomly sheared using a nebulizer and the 1 to 2 Kb range was gel purified and cloned into the dephosphoryated HincII site of the vector"

ORIGIN

Query Match	70.4%;	Score 20;	DB 9;	Length 401;
Best Local Similarity	58.3%;	Pred. No. 1.6e+02;		
Matches	14;	Conservative	9;	Mismatches 1; Indels 0; Gaps 0;

QY 3 UNUUCGACGAGUACUCDTDCDA 26
 :::|:||||:|:|:|:|:|:|:|:

Db 152 TTATTCGGCAGCTACTCTTCAA 175

RESULT 4

AQ444240

LOCUS

DEFINITION

GSSTc0285 Trypanosoma cruzi random genomic library Trypanosoma cruzi genomic clone G11B6, genomic survey sequence.

ACCSSION

AQ444240

VERSION

AQ444240.3 GI:10130805

KEYWORDS

GSS.

SOURCE

Trypanosoma cruzi

ORGANISM

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma; Schizotrypanum.

REFERENCE

1 (bases 1 to 411)

AUTHORS

Agudo,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.

TITLE

A random sequencing approach for the analysis of the trypanosoma cruzi genome: general structure, large gene and repetitive DNA families, and gene discovery

JOURNAL

Genome Res. 10 (12), 1996-2005 (2000)

FURNED

11116094

COMMENT

On Sep 14, 2000 this sequence version replaced gi:9372589.

Contact: Sanchez D.O.
Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral San Martin)
Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos Aires, Argentina
Tel: (54-11) 4580/7255/7
Fax: (54-11) 4752-9639
Email: dsanchez@ib.unsam.edu.ar
Sequences were basecalled with phred and vector was masked with crossmatch (see http://genome.washington.edu). Sequences were then trimmed from both ends to remove low quality bases and masked vector.

Seq primer: T7
Class: shotgun.

FEATURES

source	Location/Qualifiers
	1..411
	/organism="Trypanosoma cruzi"
	/mol_type="genomic DNA"
	/strain="CL-Brener"

69.0%: score 19.6; DB 2: Length 383.

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 648)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Ceiniker, S. and Rubin, G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11586823.
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003509: arm.X [18190295,18494054]
estimated-cyto:17B1-17D3: 04/07/2001
Plate: AT.265 row: D column: 3
High quality sequence stop: 552.
High quality sequence stop: 550.

FEATURES
source
1. 560
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT26539"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 648)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Ceiniker, S. and Rubin, G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11573523.
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003509: arm.X [18190295,18494054]
estimated-cyto:17B1-17D3: 04/09/2001
Plate: AT.265 row: D column: 3
High quality sequence stop: 552.
High quality sequence stop: 550.

FEATURES
source
1. 560
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT26539"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 648)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Ceiniker, S. and Rubin, G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003509: arm.X [18190295,18494054]
estimated-cyto:17B1-17D3: 04/09/2001
Plate: AT.265 row: D column: 3
High quality sequence stop: 552.
High quality sequence stop: 550.

FEATURES
source
1. 560
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT26539"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 648)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Ceiniker, S. and Rubin, G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11573523.
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003509: arm.X [18190295,18494054]
estimated-cyto:17B1-17D3: 04/09/2001
Plate: AT.265 row: D column: 3
High quality sequence stop: 552.
High quality sequence stop: 550.

FEATURES
source
1. 560
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT26539"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 648)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Ceiniker, S. and Rubin, G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11573523.
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003509: arm.X [18190295,18494054]
estimated-cyto:17B1-17D3: 04/09/2001
Plate: AT.265 row: D column: 3
High quality sequence stop: 552.
High quality sequence stop: 550.

FEATURES
source
1. 560
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT26539"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 648)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Ceiniker, S. and Rubin, G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003509: arm.X [18190295,18494054]
estimated-cyto:17B1-17D3: 04/09/2001
Plate: AT.265 row: D column: 3
High quality sequence stop: 552.
High quality sequence stop: 550.

FEATURES
source
1. 560
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT26539"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

COMMENT

On Dec 6, 2000 this sequence version replaced gi:11582735.

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, estfruitfly.berkeley.edu

hit genomic AB003509: arm:X [18190295,18494054]

estimated-cyto:17B1-17D3: 04/09/2001

Plate: AT.139 row: H column: 5

High quality sequence stop: 582.

Location/Qualifiers

1..653

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="AT13989"

/sex="male"

/dev_stages="0-3 day old Ore-R males"

/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates

AT.121-AT.319: DHS-alpha TONA"

/clone_lib="AT Drosophila melanogaster adult testes pOTB7"

/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;

Site 2: XhoI; The mRNA for the testis library was made

from testes and seminal vesicles hand dissected from 0-3

day old Ore-R males. RNA kindly provided by the lab of

Margaret Fuller. Sized fractionated cDNAs were directly

ligated into pOTB7. Plasmid cDNA library."

ORIGIN

Query Match

Best Local Similarity 69.0%; Score 19.6; DB 2; Length 653;

Matches 18; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy

1 CGUAUUCGACGACGACUCUCUUCDCDADGDCDCDT 34

Db

327 CGGATTCTCAGCAGCAGTCATCGCGCGTCGCAT 294

RESULT 10

LOCUS

BF486138 703 bp mRNA linear EST 19-APR-2001

AT20339.5prime AT Drosophila melanogaster adult testes pOTB7

Drosophila melanogaster cDNA clone AT20339 5 similar to CG6481:

FBan006481 located on: X 17B5-17B5; 04/09/2001, mRNA sequence.

ACCESSION

BF486138

VERSION

BF486138.2 GI:13698027

KEYWORDS

EST.

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 703)

REFERENCE

AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,

Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,

Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,

Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,

Pacaleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,

Lewis, S.E., Celniker, S. and Rubin, G.M.

BDGP/HMI AT Drosophila EST Project

Unpublished (2000)

TITLE

JOURNAL

COMMENT

On Dec 6, 2000 this sequence version replaced gi:11569439.

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, estfruitfly.berkeley.edu

hit genomic AB003509: arm:X [18190295,18494054]

estimated-cyto:17B1-17D3: 04/09/2001

Plate: AT.203 row: D column: 3

High quality sequence stop: 702.

FEATURES

source

Location/Qualifiers

1..703

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="AT20339"

/sex="male"

/dev_stages="0-3 day old Ore-R males"

/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates

AT.121-AT.319: DHS-alpha TONA"

/clone_lib="AT Drosophila melanogaster adult testes pOTB7"

/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;

Site 2: XhoI; The mRNA for the testis library was made

from testes and seminal vesicles hand dissected from 0-3

day old Ore-R males. RNA kindly provided by the lab of

Margaret Fuller. Sized fractionated cDNAs were directly

ligated into pOTB7. Plasmid cDNA library."

ORIGIN

Query Match

Best Local Similarity 69.0%; Score 19.6; DB 2; Length 703;

Matches 18; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy

1 CGUAUUCGACGACGACUCUCUUCDCDADGDCDCDT 34

Db

73 CGGATTCTCAGCAGCAGTCATCGCGCGTCGCAT 40

RESULT 11

EX653933/c

LOCUS

DEFINITION

Arabidopsis thaliana T-DNA flanking sequence GK-593C03-021827,

genomic survey sequence.

ACCESSION

EX653933

VERSION

EX653933.1 GI:37610321

KEYWORDS

GSS.

SOURCE

ORGANISM

Arabidopsis thaliana (chale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1

Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P., and Weissshaar, B.

GABI-Kat Simplesearch: a flanking sequence tag (FST) database for

the identification of T-DNA insertion mutants in Arabidopsis

thaliana

Bioinformatics 19 (11), 1441-1442 (2003)

2

ROSSO, M.G., LI, Y., STRIZHOV, N., REISS, B., DEKKER, K. and

WEISSSHAAR, B.

An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for

flanking sequence tag-based reverse genetics

Plant Mol. Biol. 53 (1-2), 247-259 (2003)

3

Strizhov, N., Li, Y., Rosso, M.G., Viehoever, P., Dekker, K.A. and

Weissshaar, B.

High-throughput generation of sequence indexes from T-DNA

mutagenized Arabidopsis thaliana lines

Biotechniques 35 (6), 1164-1168 (2003)

4

(bases 1 to 168)

Strizhov, N., Li, Y., Rosso, M.G. and Weissshaar, B.

Direct Submission

Submitted (31-MAR-2004) Weissshaar B., Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

This sequence has been recovered from the left border of the T-DNA.

It indicates an insertion close to or within gene At2g05420.

Details on the protocols used for generation of the sequence are

described in References 1-3. The sequences are generated at the MPI

for Plant Breeding Research in the context of the GABI-Kat project.

GABI-Kat is part of the German Plant Genomics program designated

COMMENT

REFERENCE

1 (bases 1 to 444)
 Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
 Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 CONTACT: Nahm B.H.
 Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
 of BioScience and Bioinformatics, Myongui University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.

FEATURES

source

1..444
 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:39947"
 /clone="ABP-06-K22"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"
 /clone_lib="ABP3-overexpressing transgenic rice plasmid
 cDNA library (ABP)"
 /notes="vector: PCR4-TOPO; Site_1: EcoRI; Leaf was dried
 for 2hrs. Oligo-capped mRNA was reverse transcribed and
 then used for PCR. mRNA was prepared from ABA-responsive
 element binding transcription factor 3 overexpression
 line."

ORIGIN

Query Match 68.3%; Score 19.4; DB 6; Length 444;
 Best Local Similarity 48.5%; Pred. NO. 3.1e+02;
 Matches 16; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
 QY 2 GUUUCUGACGACGUACUCDTDCDADGDCDT 34
 DB 56 GGATTCGACGTGTACTGATCAAGTCGGTT 24

RESULT 15

CO335177 477 bp mRNA linear EST 29-JUN-2004
 LOCUS EK316044.5prime Exelixis Flytag CK01 pCDNA-SK+ Drosophila
 melanogaster cDNA clone EK316044 5, mRNA sequence.
 CO335177
 ACCESSION CO335177.1 GI:49395452
 VERSION
 KEYWORDS
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 477)
 Kopczynski, C., Platt, D., Campbell, J., Muzong, C., Laufer, A.,
 Peterson, E. and Swimmer, C.
 Exelixis Flytag EST Project CK01 Library
 Unpublished (2004)
 CONTACT: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Plate: EK.3160 row: D column: 8
 High quality sequence stop: 393.

FEATURES

source

1..477
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="EK316044"
 /clone_lib="Exelixis Flytag CK01 pCDNA-SK+"

/note="Organ: mixed stage embryos, imaginal disks, and
 adult heads; Vector: pCDNA-SK+; Site_1: NotI; Site_2:
 XhoI; Random primed, normalized library from mixed stage
 embryos, imaginal disks, and adult heads."

ORIGIN

Query Match 68.3%; Score 19.4; DB 7; Length 477;
 Best Local Similarity 51.7%; Pred. NO. 3.2e+02;
 Matches 15; Conservative 11; Mismatches 3; Indels 0; Gaps 0;
 QY 3 UAUUCUGACGACGUACUCDTDCDADGDCD 31
 DB 356 TATTCAGACGACGTCTCTTCTTGAGACA 384

Search completed: February 18, 2006, 16:24:08
 Job time : 2105.52 secs

this Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2006, 15:01:08 ; Search time 118.452 Seconds
(without alignments)
510.226 Million cell updates/sec

Title: US-10-824-829-2

Perfect score: 28.4

Sequence: 1 cgaucgacgacguacucdtdcdgdcdcdt 34

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA.*
- 1: /cgm2_6/ptodata/1/ina/1 COMB.seq.*
 - 2: /cgm2_6/ptodata/1/ina/5 COMB.seq.*
 - 3: /cgm2_6/ptodata/1/ina/6A COMB.seq.*
 - 4: /cgm2_6/ptodata/1/ina/6B COMB.seq.*
 - 5: /cgm2_6/ptodata/1/ina/H COMB.seq.*
 - 6: /cgm2_6/ptodata/1/ina/PC/US COMB.seq.*
 - 7: /cgm2_6/ptodata/1/ina/PP COMB.seq.*
 - 8: /cgm2_6/ptodata/1/ina/RE COMB.seq.*
 - 9: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.6	65.5	42988	3	US-08-311-731A-128
2	18	63.4	1419	3	Sequence 128, App
3	18	63.4	53500	3	Sequence 61, Appl
4	17.6	62.0	5840	3	Sequence 76, Appl
5	17.6	62.0	312474	3	Sequence 5692, Ap
6	17.4	61.3	440	3	Sequence 17434, A
7	17.4	61.3	1154	3	Sequence 751, App
8	17.2	60.6	9871	3	Sequence 43, Appl
9	17.2	60.6	9871	3	Sequence 24, Appl
10	17.2	60.6	9871	3	Sequence 24, Appl
11	17.2	60.6	9871	3	Sequence 24, Appl
12	17.2	60.6	9871	3	Sequence 24, Appl
13	17.2	60.6	9871	3	Sequence 24, Appl
14	17.2	60.6	9871	3	Sequence 24, Appl
15	17.2	60.6	9871	3	Sequence 24, Appl
16	17.2	60.6	9871	3	Sequence 24, Appl
17	17	59.9	168	3	Sequence 2505, Ap
18	17	59.9	729	3	Sequence 2082, Ap
19	17	59.9	10060	3	Sequence 25, Appl
20	17	59.9	10060	3	Sequence 25, Appl
21	17	59.9	10060	3	Sequence 25, Appl
22	17	59.9	10060	3	Sequence 25, Appl
23	17	59.9	10060	3	Sequence 25, Appl
24	17	59.9	10060	3	Sequence 25, Appl

25	17	59.9	10060	3	US-09-479-123-25	Sequence 25, Appl
26	17	59.9	10060	3	US-09-484-317A-25	Sequence 25, Appl
27	17	59.9	10060	3	US-09-276-820A-25	Sequence 25, Appl
28	17	59.9	14367	3	US-09-902-540-1113	Sequence 1113, Ap
29	17	59.9	234884	3	US-09-949-016-16420	Sequence 16420, A
30	17	59.9	678533	3	US-09-949-016-14577	Sequence 14577, A
31	17	59.9	678533	3	US-09-949-016-14578	Sequence 14578, A
32	16.8	59.2	270	3	US-09-902-540-1742	Sequence 9415, Ap
33	16.8	59.2	272	3	US-09-902-540-1742	Sequence 1742, Ap
34	16.8	59.2	512	2	US-08-867-820A-8	Sequence 8, Appl
35	16.8	59.2	513	2	US-08-867-820A-7	Sequence 7, Appl
36	16.8	59.2	513	2	US-08-867-820A-20	Sequence 2, Appl
37	16.8	59.2	514	2	US-08-867-820A-2	Sequence 2, Appl
38	16.8	59.2	514	2	US-08-867-820A-3	Sequence 3, Appl
39	16.8	59.2	514	2	US-08-867-820A-4	Sequence 4, Appl
40	16.8	59.2	514	2	US-08-867-820A-6	Sequence 6, Appl
41	16.8	59.2	514	2	US-08-867-820A-21	Sequence 21, Appl
42	16.8	59.2	514	2	US-08-867-820A-22	Sequence 22, Appl
43	16.8	59.2	514	2	US-08-867-820A-28	Sequence 28, Appl
44	16.8	59.2	514	2	US-08-867-820A-30	Sequence 30, Appl
45	16.8	59.2	514	2	US-08-867-820A-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1

US-08-311-731A-128/c
; Sequence 128, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42988 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-128

Query Match 65.5%; Score 18.6; DB 3; Length 42988;
Best Local Similarity 48.5%; Pred. No. 62;
Matches 16; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 2 GUUAUCGACGACGUACUCUDDCDADGDCDDT 34
Db 8776 GTGTTCTGTTGAAGCACTCTTTGAGAGCGCCT 8744

RESULT 2
US-09-266-965-61/c
; Sequence 61, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456U51
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 1419
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-266-965-61

Query Match 63.4%; Score 18; DB 3; Length 1419;
Best Local Similarity 55.9%; Pred. No. 51;
Matches 19; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGUAUCGACGACGUACUCUDDCDADGDCDDT 34
Db 1241 GCGATGCTGCGACGAACTCTGCTGCGCGCGT 1208

RESULT 3
US-09-266-965-76
; Sequence 76, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456U51
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 76
; LENGTH: 53500
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-266-965-76

Query Match 63.4%; Score 18; DB 3; Length 53500;
Best Local Similarity 55.9%; Pred. No. 1.3e+02;
Matches 19; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGUAUCGACGACGUACUCUDDCDADGDCDDT 34
Db 35459 GCGATGCTGCGACGAACTCTGCTGCGCGCGT 35492

RESULT 4
US-09-949-016-5692/c
; Sequence 5692, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5692
; LENGTH: 5840
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5692

Query Match 62.0%; Score 17.6; DB 3; Length 5840;
Best Local Similarity 46.4%; Pred. No. 1.2e+02;
Matches 13; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

QY 2 GUUAUCGACGACGUACUCUDDCDADGD 29
Db 5109 GTATTCTGATGAGTACTCTTTAAGCTGT 5082

RESULT 5
US-09-949-016-17434/c
; Sequence 17434, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17434
; LENGTH: 312474
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17434

Query Match 62.0%; Score 17.6; DB 3; Length 312474;
Best Local Similarity 46.4%; Pred. No. 3.1e+02;
Matches 13; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GUUUUCUGACGACGUACUCUDTDCDADGD 29
DB 309742 GTATTCTGATGAGGTACTCTTAAGCTGT 309715

RESULT 6
US-09-533-559-751
; Sequence 751, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 751
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(440)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-751

Query Match 61.3%; Score 17.4; DB 3; Length 440;
Best Local Similarity 48.1%; Pred. No. 75;
Matches 13; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

Qy 3 UAUUCUGACGACGUACUCUDTDCDADGD 29
DB 171 TATTCTGTCGAGGTACTCGAGCCATGG 197

RESULT 7
US-09-273-164-43/c
; Sequence 43, Application US/09273164
; Patent No. 6869770
; GENERAL INFORMATION:
; APPLICANT: Roberts, R. Michael
; APPLICANT: Green, Jonathan
; APPLICANT: Xie, Sancel
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR EARLY PREGNANCY DIAGNOSIS
; FILE REFERENCE: UVM003/UVM003P
; CURRENT APPLICATION NUMBER: US/09/273,164
; PRIOR FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 60/078,783
; PRIOR FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: 60/106,188
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1154
; TYPE: DNA
; ORGANISM: bovidae
US-09-273-164-43

Query Match 61.3%; Score 17.4; DB 3; Length 1154;
Best Local Similarity 48.4%; Pred. No. 96;
Matches 15; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GUUUUCUGACGACGUACUCUDTDCDADGD 32

DB 69 GTATTTCGACGACCTCTGAGAGGCCAC 39

RESULT 8
US-09-479-122-24
; Sequence 24, Application US/09479122
; Patent No. 6410266
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003C
; CURRENT APPLICATION NUMBER: US/09/479,122
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8481)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8633)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-122-24

Query Match 60.6%; Score 17.2; DB 3; Length 9871;
Best Local Similarity 46.7%; Pred. No. 2.1e+02;
Matches 14; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GUUUUCUGACGACGUACUCUDTDCDADGD 31
DB 759 GAATTCGACGACCTCTGATTAAGATCT 788

RESULT 9
US-09-484-997-24
; Sequence 24, Application US/09484997
; Patent No. 6524818
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003J
; CURRENT APPLICATION NUMBER: US/09/484,997
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022

```

; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8481)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8633)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-997-24

Query Match      60.6%; Score 17.2; DB 3; Length 9871;
Best Local Similarity 46.7%; Pred. No. 2.1e+02;
Matches 14; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY      2 GUUUUCGACGACGUACUCDTDCDADGDCD 31
Db      759 GAATTCGACGACCTACTGATTAAAGATCT 788

RESULT 10
US-09-481-355-24
; Sequence 24, Application US/09481355
; Patent No. 6524824
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003F
; CURRENT APPLICATION NUMBER: US/09/481,355
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8481)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8633)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-481-355-24

Query Match      60.6%; Score 17.2; DB 3; Length 9871;
Best Local Similarity 46.7%; Pred. No. 2.1e+02;
Matches 14; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY      2 GUUUUCGACGACGUACUCDTDCDADGDCD 31
Db      759 GAATTCGACGACCTACTGATTAAAGATCT 788

RESULT 11
US-09-481-355-24
; Sequence 24, Application US/09481355
; Patent No. 6524824
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003F
; CURRENT APPLICATION NUMBER: US/09/481,355
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8481)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8633)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-481-355-24

Query Match      60.6%; Score 17.2; DB 3; Length 9871;
Best Local Similarity 46.7%; Pred. No. 2.1e+02;
Matches 14; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY      2 GUUUUCGACGACGUACUCDTDCDADGDCD 31
Db      759 GAATTCGACGACCTACTGATTAAAGATCT 788

RESULT 12
US-09-455-659A-24
; Sequence 24, Application US/09455659A
; Patent No. 6602686
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA

```

```

US-09-481-282-24
; Sequence 24, Application US/09481282
; Patent No. 6541221
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003US
; CURRENT APPLICATION NUMBER: US/09/481,282
; CURRENT FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8481)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8633)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-481-282-24

Query Match      60.6%; Score 17.2; DB 3; Length 9871;
Best Local Similarity 46.7%; Pred. No. 2.1e+02;
Matches 14; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY      2 GUUUUCGACGACGUACUCDTDCDADGDCD 31
Db      759 GAATTCGACGACCTACTGATTAAAGATCT 788

RESULT 12
US-09-455-659A-24
; Sequence 24, Application US/09455659A
; Patent No. 6602686
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA

```

```

* APPLICANT: SHERP, BRUCE
* APPLICANT: RUNDLETT, STEPHEN
* TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
* TITLE OF INVENTION: ENDOGENOUS GENES
* FILE REFERENCE: 0221-0003B
* CURRENT APPLICATION NUMBER: US/09/479,123
* CURRENT FILING DATE: 2000-01-07
* PRIOR APPLICATION NUMBER: 09/263,814
* PRIOR FILING DATE: 1999-03-08
* PRIOR APPLICATION NUMBER: 09/253,022
* PRIOR FILING DATE: 1999-02-19
* PRIOR APPLICATION NUMBER: 09/159,643
* PRIOR FILING DATE: 1998-09-24
* PRIOR APPLICATION NUMBER: 08/941,223
* PRIOR FILING DATE: 1997-09-26
* NUMBER OF SEQ ID NOS: 33
* SOFTWARE: PatentIn Ver. 2.1
* SEQ ID NO 24
* LENGTH: 9871
* TYPE: DNA
* ORGANISM: Homo sapiens
* FEATURE:
* NAME/KEY: modified_base
* LOCATION: (8481)
* OTHER INFORMATION: a, c, t, g, other or unknown
* FEATURE:
* NAME/KEY: modified_base
* LOCATION: (8633)
* OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-123-24

Query Match          60.6%; Score 17.2; DB 3; Length 9871;
Best Local Similarity 46.7%; Pred. No. 2.1e+02;
Matches 14; Conservative 11; Mismatches 5; Indels 0; Gaps 0

Qy      2  GUUUUGAGCGACGUACUCDTDCDADGDCD 31
Db      | : : : : : : : : : : : : : : : : : :
       759 GAATTCGACGACCTACTGATTAAAGATCT 788

RESULT 15
US-09-484-317A-24
* Sequence 24, Application US/09484317A
* Patent No. 6740503
* GENERAL INFORMATION:
* APPLICANT: HARRINGTON, JOHN J.
* APPLICANT: SHERP, BRUCE
* APPLICANT: RUNDLETT, STEPHEN
* TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
* TITLE OF INVENTION: ENDOGENOUS GENES
* FILE REFERENCE: ATX-007CP4DV11
* CURRENT APPLICATION NUMBER: US/09/484,317A
* CURRENT FILING DATE: 2000-01-18
* PRIOR APPLICATION NUMBER: 09/276,820
* PRIOR FILING DATE: 1999-03-26
* PRIOR APPLICATION NUMBER: 09/263,814
* PRIOR FILING DATE: 1999-03-08
* PRIOR APPLICATION NUMBER: 09/253,022
* PRIOR FILING DATE: 1999-02-19
* PRIOR APPLICATION NUMBER: 09/159,643
* PRIOR FILING DATE: 1998-09-24
* PRIOR APPLICATION NUMBER: 08/941,223
* PRIOR FILING DATE: 1997-09-26
* NUMBER OF SEQ ID NOS: 33
* SOFTWARE: PatentIn Ver. 2.1
* SEQ ID NO 24
* LENGTH: 9871
* TYPE: DNA
* ORGANISM: Homo sapiens
* FEATURE:
* NAME/KEY: modified_base
* LOCATION: (8481)
* OTHER INFORMATION: a, c, t, g, other or unknown

```

```
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8633)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-317A-24
```

```
Query Match      60.6%; Score 17.2; DB 3; Length 9871;
Best Local Similarity 46.7%; Pred. No. 2.1e+02;
Matches 14; Conservative 11; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy      2  GUUUUCUGACGACGUACUCUDTDCDADGDCD 31
Db      759  GAATTCTGACGACCTACTGATTAAGATCT 788
```

```
Search completed: February 18, 2006, 15:20:41
Job time : 120.452 secs
```